

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2004, 12:22:29 ; Search time 826 Seconds
(without alignments)
10524.263 Million cell updates/sec

Title: US-10-698-235-1
Perfect score: 1656
Sequence: 1 aagtaataatcacatagagg.....gggtgtkgaaraataaacg 1656

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	292.4	17.7	110000	2 AAT42063_09	Continuation (10 o
2	197.8	11.9	1572	3 AAAL5298	AAAL5298 DNA encod
3	196.2	11.8	1575	4 AAAL12105	AAAL12105 Neisseria
4	196.2	11.8	12438	3 AAAL1550	AAAL1550 N. mening
5	196.2	11.8	110000	3 AAAL1489_6	Continuation (7 of
6	196.2	11.8	172325	3 AAAL1613	AAAL1613 Neisseria
7	196.2	11.8	349980	3 AAAL1612	AAAL1612 Neisseria
8	185.8	11.2	1467	10 ABZ41628	ABZ41628 N. gonorr
9	185.8	11.2	1575	2 AAAL12107	AAAL12107 Neisseria
10	185.8	11.2	1575	3 AAAL15323	AAAL15323 DNA encod
11	120.8	7.3	1275	2 AAAL12106	AAAL12106 Neisseria
12	96.6	5.8	110000	10 ACF65385_0	ACF65385 Photorhab
13	96.6	5.8	110000	10 ACF67367_29	Continuation (30 o
14	95.4	5.8	1284	10 ACF69918	ACF69918 Photorhab
15	95.4	5.8	1335	9 ADA31630	ADA31630 DNA encod
16	91.4	5.5	47108	6 ABK31510	ABK31510 Signal tr
17	89.4	5.4	8056	8 ABZ10246	ABZ10246 Haematopo
18	88.2	5.3	1377	11 ABD11809	ABD11809 Pseudomon
19	88.2	5.3	1383	11 ABD11702	ABD11702 Pseudomon
20	88.2	5.3	1410	11 ABD12145	ABD12145 Pseudomon
21	87.2	5.3	19380	6 AA561427	AA561427 Human gen

ALIGNMENTS

RESULT 1

AAT42063_09

Continuation (10 of 19) of AAT42063 from base 900001 (Haemophilus influenzae complete ge

WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

WP	Fragment Name	Begin	End
WP	AAT42063_00	1	110000
WP	AAT42063_01	100001	210000
WP	AAT42063_02	200001	310000
WP	AAT42063_03	300001	410000
WP	AAT42063_04	400001	510000
WP	AAT42063_05	500001	610000
WP	AAT42063_06	600001	710000
WP	AAT42063_07	700001	810000
WP	AAT42063_08	800001	910000
WP	AAT42063_09	900001	1010000
WP	AAT42063_10	1000001	1110000
WP	AAT42063_11	1100001	1210000
WP	AAT42063_12	1200001	1310000
WP	AAT42063_13	1300001	1410000
WP	AAT42063_14	1400001	1510000
WP	AAT42063_15	1500001	1610000
WP	AAT42063_16	1600001	1710000
WP	AAT42063_17	1700001	1810000
WP	AAT42063_18	1800001	1830121

Query Match 17.7%; Score 292.4; DB 2; Length 110000;

Best Local Similarity 85.1%; Pred. No. 3.3e-43;

Matches 336; Conservative 1; Mismatches 57; Indels 1; Gaps 1;

QY	8	ATCACAATAGSGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCAGAGGGATATCC	67
DB	41919	ATATCACAAATGGATTCGCGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGGATATCC	41977
QY	68	AAAGGATTATCCAAAGGCGATGAATCAGCGGATTGCCCTAAAGTTCGGTCAAAAACGA	127
DB	41978	AAAGATTATCGAAAGGCGATGAATCAGCGGATTGCCCTAAAGTTCGGTCAAAAACGA	42037
QY	128	GAAGATTTCTTAGCGGGTGTGCGAGAACAGAGCAAGAGTGTAGTACAAAACGGCGGTG	187
DB	42038	GAAGATTTCTTAGCGGGTGTGCGAGAACAGAGCAAGAGTGTAGTACAAAACGGCGGTG	42097
QY	188	TGTACTTTGTGTGACTCGGTAGGCGAAAGTGTATTGAGGACACAAAGACGTTAA	247
DB	42098	TGTACTTTGTGTGACTCGGTAGGCGAAAGTGTATTGAGGACACAAAGACGTTAA	42157

XX	08-OCT-1999	(first entry)	
XX	DE	Neisseria meningitidis complete ORF81 sequence.	
XX	KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;	
XX	KW	treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.	
XX	OS	Neisseria meningitidis.	
XX	XX	WO9924578-A2.	
XX	PN	20-MAY-1999.	
XX	PD		
XX	PF	09-OCT-1998; 98WO-IB001665.	
XX	PR	06-NOV-1997; 97GB-00023516.	
XX	PR	14-NOV-1997; 97GB-00024190.	
XX	PR	18-NOV-1997; 97GB-00024386.	
XX	PR	27-NOV-1997; 97GB-00025158.	
XX	PR	10-DEC-1997; 97GB-00026147.	
XX	PR	14-JAN-1998; 98GB-00000759.	
XX	PR	01-SEP-1998; 98GB-00019016.	
XX	PA	(CHIR-) CHIRON SPA.	
XX	PI	Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;	
XX	XX	WPI; 1999-327407/27.	
XX	DR	P-PSDB; AAY38651.	
XX	XX		
XX	PT	Proteins from Neisseria meningitidis and N. gonorrhoeae useful for	
XX	PT	diagnosis, treatment and prevention of infection.	
XX	XX		
XX	PS	Claim 9; Page 205; 524pp; English.	
XX	CC	Nucleotide sequences AAZ11972-Z12358 represent open reading frames (ORFs)	
XX	CC	of Neisseria meningitidis and N. gonorrhoeae which encode antigenic	
XX	CC	proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,	
XX	CC	their nucleic acids and antibodies are used for diagnosis, prevention (as	
XX	CC	vaccines) or treatment of Neisseria infections, such as meningitis,	
XX	CC	septicaemia and gonorrhea. Both organisms are closely related. Fragments	
XX	CC	of the nucleic acids are useful as hybridisation probes and antisense	
XX	CC	reagents	
XX	XX		
XX	SQ	Sequence 1575 BP; 376 A; 378 C; 405 G; 416 T; 0 U; 0 Other;	
	Query Match	11.8%; Score 196.2; DB 2; Length 1575;	
	Best Local Similarity	51.5%; Pred. No. 4.6e-26;	
	Matches 532; Conservative 2; Mismatches 481; Indels 18; Gaps 3;		
Qy	589	CTGAAGCTATTTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTTCAAAATATAGAA	648
Db	110	CGGAAACGTTGCGCTGACATTTGTGATTGCTGCGCTGATCTGTTGCGGTTATAAGG	169
Qy	649	TTACAAGAAATATTAGTTGGCGCTTTATTTCGAATAAGATTGTTGTTTAATAATGTACATT	708
Db	170	TGACCGCTTTGTTGATTGCGGTGTTTTTGGCTTCAGCAATTATGCCAACATGTGCATT	229
Qy	709	ATGCAGTATACCAATCTTGGATTGGACCTGTTAAATTACTCTGCAATTTAAAGAAATTA	768
Db	230	ACGCGTTTATCAAAGCTGGATGACGGGCATCAATTATTGCTGATGCTGAAGAGGTTA	289
Qy	769	ATGAGATAACAATGCTGGCTTAACAATGATAGATAAATCATATATCCATTGTTATTG	828
Db	290	CCGAAGTCGGCAGCGCGGTGCGTCGATGTTGGATAAGTTGGGCTGCTGTTGTGCGG	349
Qy	829	GTTTATTGGAAGTGTCTGTGTTTTTAAGTTTAAAGTTTTCATAAAAGAAAGTATATAAC	888
Db	350	CGGTGTTGGAAGTCATGTTGTTTTTGACCTTGCCAAAGTTCGCCGCTAAGACGCA-T	405
Qy	889	TTTCTTGGAATTTTGGACTTTTATTTTTTATGCTGTGATGATGATGTTTTTGTTCGAGCGT	948

Qy	1426	TGCATATGATCAATTTT	TAGGAGGAGCTTGG	ATATGATATCGT	TTTTCCAGATAATG	1485
Db	935	TGGCAATTTTGA	CTTAATCGT	TAAGAAATGG	ATAGACCATCTG	994
Qy	1486	AAGGGTATCTTT	TAAGAGATTCA	ATGCTCGT	GATATAAATTAC	1545
Db	995	TTGGCT	ACGGCAACGG	CGACATATG	CCGATGAGAGCTG	1054
Qy	1546	TTAATTT	PAGATAATG	GGTTATCAT	TTTGTGTTTACAT	1605
Db	1055	TCAATTT	GGCAGCGG	CAGGCA	TTTATTCGTGTTG	1114
Qy	1606	ATGGGC	CATTATT	1618		
Db	1115	ACGGCG	CATTGTT	1127		
RESULT 10						
AA15323						
ID	AA15323 standard; DNA; 1575 BP.					
XX	AA15323;					
XX	15-SEP-2003 (revised)					
DT	04-SEP-2000 (first entry)					
XX	DNA encoding a polypeptide of a <i>Neisseria</i> pathogenic strain.					
DB	Pathogenic strain; <i>Neisseria</i> ; vaccine; <i>Neisseria</i> infection; ss.					
KW	<i>Neisseria</i> gonorrhoeae.					
OS						
XX						
FH	Key	Location/Qualifiers				
FT	CDS	1..1575				
FT		/*tag= a				
XX	WO200026375-A2.					
PN						
XX	11-MAY-2000.					
PD						
XX	28-OCT-1999; 99WO-FR002643.					
PF						
XX	30-OCT-1998; 98FR-00013693.					
PR						
XX	(INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.					
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.					
XX						
PI	Aujaume L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;					
PI	Tinsley C, Perrin A;					
XX						
DR	WPI; 2000-365622/31.					
DR	P-PSDB; AAY93293.					
XX	New polypeptide specific for pathogenic <i>Neisseria</i> useful in therapeutic					
PT	or preventative vaccines and for diagnosis.					
XX						
PS	Claim 4; Page 137-139; 187pp; French.					
XX						
CC	The present sequence encodes a protein that is specific for pathogenic					
CC	strains of <i>Neisseria</i> . The polynucleotides, polypeptides, or their					
CC	antigenic fragments, are used in vaccines to treat or protect against					
CC	<i>Neisseria</i> infections, particularly by <i>N. meningitidis</i> . The polynucleotide					
CC	sequence is also used for recombinant production of the polypeptide and					
CC	to produce attenuated <i>Neisseria</i> strains that overexpress it, or express					
CC	it in a non-toxic mutant form. (Updated on 15-SEP-2003 to standard OS					
CC	field)					
XX						
SQ	Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 U; 0 Other;					
Query Match 11.2%; Score 185.8; DB 3; Length 1575;						
Best Local Similarity 50.7%; Pred. No. 3.4e-24;						
Matches 524; Conservative 4; Mismatches 487; Indels 18; Gaps 3;						

[illegible]

Query Match 11.2%; Score 185.8; DB 3; Length 1575;
Best Local Similarity 50.7%; Pred. No. 3.4e-24;
Matches 524; Conservative 4; Mismatches 487; Indels 18;

PI	Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A; Buchrieser C;	WP	ACF67367_09	900001	1010000
PI		WP	ACF67367_10	1000001	1110000
XX	WPI; 2003-148459/14.	WP	ACF67367_11	1100001	1210000
DR		WP	ACF67367_12	1200001	1310000
XX		WP	ACF67367_13	1300001	1410000
PT	Genomic sequence of Photorhabdus luminescens and encoded polypeptides, useful e.g. as therapeutic antimicrobials and agricultural pesticides.	WP	ACF67367_14	1400001	1510000
XX		WP	ACF67367_15	1500001	1610000
PS	Claim 1; SEQ ID NO 38; 1205pp; French.	WP	ACF67367_16	1600001	1710000
XX		WP	ACF67367_17	1700001	1810000
XX		WP	ACF67367_18	1800001	1910000
CC	The invention relates to the isolation of genes and their encoded proteins from Photorhabdus luminescens. The isolated sequences are sources of probes and primers for detecting the genome of P. luminescens and related species; to study polymorphisms; for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than P. luminescens and are able to alter response or sensitivity to toxins and antibiotics produced by P. luminescens. Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and antibiotics. The genes are useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the genes and Ab are also useful therapeutically (to treat microbial infection by bacteria or fungi that are sensitive to P. luminescens-encoded toxins or antibiotics) and as biopesticides. Other uses of the genes and the proteins are as virulence factors and for identifying targets of human diseases for which P. luminescens is a model (particularly plague and whooping cough). This sequence represents one of the isolated P. luminescens genes	WP	ACF67367_19	1900001	2010000
XX		WP	ACF67367_20	2000001	2110000
CC		WP	ACF67367_21	2100001	2210000
CC		WP	ACF67367_22	2200001	2310000
CC		WP	ACF67367_23	2300001	2410000
CC		WP	ACF67367_24	2400001	2510000
CC		WP	ACF67367_25	2500001	2610000
CC		WP	ACF67367_26	2600001	2710000
CC		WP	ACF67367_27	2700001	2810000
CC		WP	ACF67367_28	2800001	2910000
CC		WP	ACF67367_29	2900001	3010000
CC		WP	ACF67367_30	3000001	3110000
CC		WP	ACF67367_31	3100001	3210000
CC		WP	ACF67367_32	3200001	3310000
CC		WP	ACF67367_33	3300001	3410000
CC		WP	ACF67367_34	3400001	3510000
CC		WP	ACF67367_35	3500001	3610000
CC		WP	ACF67367_36	3600001	3710000
CC		WP	ACF67367_37	3700001	3810000
CC		WP	ACF67367_38	3800001	3910000
CC		WP	ACF67367_39	3900001	4010000
XX		WP	ACF67367_40	4000001	4110000
SQ	Sequence 618776 BP; 181854 A; 139249 C; 125634 G; 172037 T; 0 U; 2 Other;	WP	ACF67367_41	4100001	4210000
	Query Match 5.8%; Score 96.6; DB 10; Length 110000; Best Local Similarity 58.0%; Pred. No. 5.1e-08; Matches 171; Conservative 0; Mismatches 124; Indels 0; Gaps 0;	WP	ACF67367_42	4200001	4310000
QY	37 TAGGTATCGTATGGCTGCAGAGGGATATCCAAAGGATTATCCAAAGCGGATGAATCA 96	WP	ACF67367_43	4300001	4410000
DB	65180 TGGGTGTGTTACTCGTGCAGAGGGTATCCTGCTGTTATATCAAAAGGGATATCATTC 65239	WP	ACF67367_44	4400001	4510000
QY	97 GCGGATTGCTAAAGTGGCGGTCAAAAACGAGAAAGTTTCTTAGCGGGTGTGCGAGAAC 156	WP	ACF67367_45	4500001	4610000
DB	65240 ATGGCTTGCCCAACAGAAATGAAGTAGCAAGGTTTCTCCATGCTGGTACAGCAATGA 65299	WP	ACF67367_46	4600001	4710000
QY	157 AAGAAGCGCACTAGTACCAAAAGCGGTCGTGTTACTTTGTGCTGCTGCTAGCGTAA 216	WP	ACF67367_47	4700001	4810000
DB	65300 AAGACAATGATGTTATTACCGCGGTGGACGCGTCTGTGTGTTACAGCATTAGGAGAA 65359	WP	ACF67367_48	4800001	4910000
QY	217 GTGTATTTGAAGCACAAACAAAGCGTTAAATTTGGCTGAGCAAAATTCATGTCGGGC 276	WP	ACF67367_49	4900001	5010000
DB	65360 CCATTATAGTGCAGAAATATGCTTACCACAGCAGAGAGATTGATGAATAACT 65419	WP	ACF67367_50	5000001	5110000
QY	277 GTTTTATCGTCAGACATTTGGTTACAGGCTGTGGAACGAGAACAAAGCAAAATA 331	WP	ACF67367_51	5100001	5210000
DB	65420 GTTTTACCGCAGGATATCGGATATCGGCGCAATTAACCGATTGAATAATCATATA 65474	WP	ACF67367_52	5200001	5310000
	RESULT 13	WP	ACF67367_53	5300001	5410000
ACF67367_29	Continuation (30 of 57) of ACF67367 from base 2900001 (Photorhabdus luminescens nucleotide sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367	WP	ACF67367_54	5400001	5510000
WP	Fragment Name Begin End	WP	ACF67367_55	5500001	5610000
WP	ACF67367_00 1 110000	WP	ACF67367_56	5600001	5648894
WP	ACF67367_01 100001 210000				
WP	ACF67367_02 200001 310000				
WP	ACF67367_03 300001 410000				
WP	ACF67367_04 400001 510000				
WP	ACF67367_05 500001 610000				
WP	ACF67367_06 600001 710000				
WP	ACF67367_07 700001 810000				
WP	ACF67367_08 800001 910000				

QY	37 TAGGTATCGTATGGCTGCAGAGGGATATCCAAAGGATTATCCAAAGCGGATGAATCA 96				
DB	70314 TGGGTGTGTTACTCGTGCAGAGGGTATCCTGCTGTTATATCAAAAGGGATATCATTC 70373				
QY	97 GCGGATTGCTAAAGTGGCGGTCAAAAACGAGAAAGTTTCTTAGCGGGTGTGCGAGAAC 156				
DB	70374 ATGGCTTGCCCAACAGAAATATGAAGTAGCAAGGTTTTCATGCTGTTACAGCAATGA 70433				
QY	157 AAGAAGCGCAAGCTAGTACAAACGCGGTCGTGTACTTTGTGTGATCTGGTTAGCGGAAA 216				
DB	70434 AAGACAATGATGTTATTACCGCGGTGGACGCGTCTGTGTGTTACAGCATTAGGAGAAA 70493				
QY	217 GTGTATTTGAAGCAACAAACAAAGCGTTAAATTTGGCTGAGCAAAATTCATGTTCTGGGC 276				
DB	70494 CCATTATAGTGCAGAAAAATGCTTACCAACAGCAGAGAGATTGAATGAATAACT 70553				
QY	277 GTTTTATCGTCAGACATTTGTTACAGGCTGTGGAACGAGAACAAAGCAAAATA 331				
DB	70554 GTTTTACCGCAGGATATCGGATATCGGCGCAATTAACCGATTGAATAATCATATA 70608				

Query Match 5.8%; Score 96.6; DB 10; Length 110000;
Best Local Similarity 58.0%; Pred. No. 5.1e-08;
Matches 171; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Db	1098	CTCAGGTATCGGTCAATCACCAGAAAGATACTAAATCTTCCATGCGAGGCACTGCTACTCG	1157
Qy	155	ACAAGAAGCAAGCTAGTCACAAACGGCGTCTGTACTTTGTGTGACTGCGTTAGGCGA	214
Db	1158	TGAAGATGGGCATATCGTTACTTCTCGCGGACGTGTACTTTGCGTGACTGCTCTAGGTGA	1217
Qy	215	AGTGTATTTGAAGCACAAACAAACCGTTAAATTTGGCTGAGCAAAATTCATGGTCTGG	274
Db	1218	TAGCGTTCTTGAAGCACAAATCAATGCTTTAGAAAGTATGTGGGCAAGTAACCTTCACAGG	1277
Qy	275	GCGTTTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAAACGAGAAACAAGCAAAATA	331
Db	1278	TATGCAATACCGCAGTGACATTGGTTACCGTGTCTATTGCTCGTGA AAAAGCTGAATA	1334

Search completed: December 23, 2004, 16:18:53
Job time : 830 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2004, 14:10:34 ; Search time 7035 Seconds
(without alignments)
11131.733 Million cell updates/sec

Title: US-10-698-235-1
Perfect score: 1656
Sequence: 1 aagtaatacacatagagg.....gggtgttgggaaraataaacg 1656

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sgs.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	1551.6	93.7 1621 1	AY599442
2	292.4	17.7 10159 1	U32770 Haemophilus
3	292.4	17.7 110000 6	Continuation (10 o
4	288.4	17.4 11122 1	AE006056 Pasteurel
5	197.8	11.9 1572 6	AX024062 Sequence
6	196.2	11.8 1575 6	A96272 Sequence 30
7	196.2	11.8 15053 1	AE002550 Neisseria
8	196.2	11.8 172325 6	AX044035 Sequence
9	196.2	11.8 349980 6	AX044034 Sequence
10	185.8	11.2 1575 6	A96276 Sequence 30
11	185.8	11.2 1575 6	AX024113 Sequence
12	134.4	8.1 288108 1	AP005083 Vibrio pa
13	121.6	7.3 300521 1	AE017153 Haemophil
14	120.8	7.3 1275 6	A96274 Sequence 30
15	120.8	7.3 349061 1	NMA222491
16	116	7.0 250150 1	AP005342
17	115.8	7.0 300169 1	AE016801 Vibrio vu
18	110	6.6 170627 2	AC125567 Rattus no
19	109	6.6 12116 1	AE004115 Vibrio ch

C 20	103	6.2 110000	1	BX950851_02	Continuation (3 of
C 21	98	5.9 188406	2	CR381547	CR381547 Danio rer
C 22	97.8	5.9 189598	2	BX957347	BX957347 Danio rer
C 23	96.6	5.8 110000	1	CR543861_24	Continuation (25 o
C 24	96.6	5.8 349287	1	BX571860	BX571860 Photorhab
C 25	96.6	5.8 349980	6	AX770907	AX770907 Sequence
C 26	95.4	5.8 1335	6	AR320367	AR320367 Sequence
C 27	95.2	5.7 192525	2	BX936428	BX936428 Danio rer
C 28	94.6	5.7 57203	3	AC115581	AC115581 Dictyoste
C 29	94.6	5.7 181381	2	CR450818	CR450818 Danio rer
C 30	94	5.7 10338	1	AE015493	AE015493 Shewanell
C 31	94	5.7 349980	6	AX344563	AX344563 Sequence
C 32	93.8	5.7 183584	9	AC012492	AC012492 Homo sapi
C 33	93.8	5.7 224322	2	CR450711	CR450711 Danio rer
C 34	92.8	5.6 250029	3	AE014839	AE014839 Plasmodiu
C 35	92.6	5.6 66993	2	AC138074	AC138074 Homo sapi
C 36	92.6	5.6 234545	5	BX470214	BX470214 Zebrafish
C 37	91.8	5.5 157591	2	BX928756	BX928756 Danio rer
C 38	91.4	5.5 47108	6	AX344506	AX344506 Sequence
C 39	91.2	5.5 67970	3	PFMAL1P3	AL031746 Plasmodiu
C 40	91.2	5.5 175282	2	CR352258	CR352258 Danio rer
C 41	90.6	5.5 903	1	AF064955	AF064955 Cxiella
C 42	90.6	5.5 187161	5	BX088526	BX088526 Zebrafish
C 43	90.4	5.5 131682	9	AL672277	AL672277 Human DNA
C 44	90.4	5.5 156975	2	CR394534	CR394534 Danio rer
C 45	90.4	5.5 157141	3	AC016445	AC016445 Drosophila

ALIGNMENTS

RESULT 1	AY599442	1621 bp	DNA	linear	BCT 09-JUN-2004
LOCUS	AY599442	1621 bp	DNA	linear	BCT 09-JUN-2004
DEFINITION	Haemophilus influenzae clone L002_HI_0151_004 putative phosphoribosylamine-glycine ligase (purD) and hypothetical protein genes, partial cds.				
ACCESSION	AY599442				
VERSION	AY599442.1	GI:48243651			
KEYWORDS	Haemophilus influenzae				
SOURCE	Haemophilus influenzae				
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.				
REFERENCE	1 (bases 1 to 1621)				
AUTHORS	Antalis,P., Shen,K., Erdos,G., Gladitz,J., Sayeed,S., Hayes,J., Ahmed,A., Johnson,S., Dice,B., Keefe,R., Dopico,R., Chong,W., Goodwin,J., Singh,M., Janto,B., Post,J.C., Ehrlich,G.D. and Hu,F.Z. Genomic sequences from a pooled library of 10 clinical isolates of Haemophilus influenzae from middle-ear effusions of pediatric patients with chronic otitis media				
TITLE	Unpublished				
JOURNAL	2 (bases 1 to 1621)				
REFERENCE	Antalis,P., Shen,K., Erdos,G., Gladitz,J., Sayeed,S., Hayes,J., Ahmed,A., Johnson,S., Dice,B., Keefe,R., Dopico,R., Chong,W., Goodwin,J., Singh,M., Janto,B., Post,J.C., Ehrlich,G.D. and Hu,F.Z. Direct Submission				
AUTHORS	Submitted (15-APR-2004) Center for Genomic Sciences, Allegheny-Singer Research Institute, 320 East North Avenue, Pittsburgh, PA 15212, USA				
TITLE	Location/Qualifiers				
JOURNAL	1. .1621				
FEATURES	/organism="Haemophilus influenzae"				
source	/mol_type="genomic DNA"				
	/db_xref="taxon:727"				
	/clone="L002_HI_0151_004"				
gene	<1. 313				
	/gene="purD"				
CDS	<1. .313				
	/gene="purD"				
	/codon_start=2				
	/transl_table=11				
	/product="putative phosphoribosylamine-glycine ligase"				

SLAFYVQGMALTYLLGLAAVAAGLPLFOIALQHPYVMVIGLSILFVALALSMFGLFTI
LPNSLQNKNTNSQKOTSGAFGAPAMGIAGVAPCTSAPLSGALLVAOSGLDLF
TGATVLLXLLGMGVPLMLITLFGNKILPKSGEMNTVKOTFEGFVMLALPVELLSRIL
PEVMEPLWAGLATVFIWALQMSKNGFGYALKIISFALAMTVQPLQNIWQOTTT
TQSAVENPVSVQKFKQIKNTEEDLRTLAENPILAMLDIYADWVACKFEKLTFS
PQVQQFNTLLQVSMTKNSPENKALMERFNVMLPTILFFDQONNEIKRSVTGFM
DADAFSNWIEKLL"
/gene="HI0886"
/note="hypothetical protein; identified by GeneMark;
putative"
/codon_start=1
/transl_table=1
/product="H. influenzae predicted coding region HI0886"
/protein_id="AAC22547.1"
/db_xref="GI:1573903"

VGDPMLLVAGVIEIVGSIILITLGLFTRQAAFILSVENAYFFLHVAGKGNLFPIAN
GGEALLYSLLFLYFVPSGACALDNKFKK"
7086..8684
/gene="HI0887"
7086..8684
/gene="HI0887"
/note="similar to SP:P15639 GB:J05126 GB:X51950 PID:147420
PID:396345 percent identity: 77.19; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=1
/product="phosphoribosylaminoimidazolecarboxamide
formyltransferase (purH)"
/protein_id="AAC22544.1"
/db_xref="GI:1573904"

/translation="MADRPTRQALLSVSDKITGIVEPAQGLVKRGVLLSTGGTAKLLA
QNALPVEVSDYTGFPENMDGRVKTLPKPHVHGILGRGTDADIMQOHLEGIDMVVV
NLVFAATVATGPOCTLDADENIDIGGFTWVRSAANKHKDVAIVANNHDFNAILAEMD
QHNSLTFETRFDLIAFHTAQYDSMIANYFQQLVKPYHIAEVEENAKCCQFPRT
LNLNFRVKRQMGSHQNAAFVYDLNVAESVATANOLOQKALSYNNIADTDADLE
CVKEFDPAACVIVKHPNCGVALGKIDLDVNRAYOTDPTSAFGGIIAFNRELDKTA
NEIVRFQFVILAPKVSAAEQVEMKKNVRLLCEGWTSSRSERLDFKRVNGGLLVQ
DADLGMVVDLLKRVQPTQELKDLFCWKVAFVKSNAIVYAKDNQITGIGAGV
MSRVYSAKIAGIKAQDGLGVACVWASDAFFPFRDIDAAAKVGICQVIHPGGSMD
QEVIDAADENHMVVLTMGRHFRH"
8780..10069
/gene="HI0888"
8780..10069
/gene="HI0888"
/note="similar to SP:P15640 GB:J05126 GB:X51950 PID:147421
PID:396344 percent identity: 75.18; identified by sequence

similarity; putative"
/codon_start=1
/transl_table=1
/product="phosphoribosylaminoimidazolecarboxamide
formyltransferase (purH)"
/protein_id="AAC22544.1"
/db_xref="GI:1573904"

/translation="MADRPTRQALLSVSDKITGIVEPAQGLVKRGVLLSTGGTAKLLA
QNALPVEVSDYTGFPENMDGRVKTLPKPHVHGILGRGTDADIMQOHLEGIDMVVV
NLVFAATVATGPOCTLDADENIDIGGFTWVRSAANKHKDVAIVANNHDFNAILAEMD
QHNSLTFETRFDLIAFHTAQYDSMIANYFQQLVKPYHIAEVEENAKCCQFPRT
LNLNFRVKRQMGSHQNAAFVYDLNVAESVATANOLOQKALSYNNIADTDADLE
CVKEFDPAACVIVKHPNCGVALGKIDLDVNRAYOTDPTSAFGGIIAFNRELDKTA
NEIVRFQFVILAPKVSAAEQVEMKKNVRLLCEGWTSSRSERLDFKRVNGGLLVQ
DADLGMVVDLLKRVQPTQELKDLFCWKVAFVKSNAIVYAKDNQITGIGAGV
MSRVYSAKIAGIKAQDGLGVACVWASDAFFPFRDIDAAAKVGICQVIHPGGSMD
QEVIDAADENHMVVLTMGRHFRH"
8780..10069
/gene="HI0888"
8780..10069
/gene="HI0888"
/note="similar to SP:P15640 GB:J05126 GB:X51950 PID:147421
PID:396344 percent identity: 75.18; identified by sequence

19985 ATTGGCTGAGCAAAATTCAATGGTCTGGGGCTTTTATCGTCGAGACATTTGTTACAGGC 10044
308 TGTGGAAACGAGCAACAGCAAAATAGTTAGAAATCTTTGTTGAATTTAATAGATAAAAAAT 367
10045 TGTGGAAACGAGAGTTCACAAAATAATCCATATCATTAATGAATTAATTTCAATTAAGT 10104
368 ATTGTACACGGGTAGAAATTTGTTATTTTCTAGGATTT 402
10105 GATGAAAATTTAGGAATGAATTTTAAATTTAAATTT 10139

RESULT 3
AR274513_09
WPCOMMENT
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name Begin End
AR274513_00 1 110000
AR274513_01 100001 210000
AR274513_02 200001 310000
AR274513_03 300001 410000
AR274513_04 400001 510000
AR274513_05 500001 610000
AR274513_06 600001 710000
AR274513_07 700001 810000
AR274513_08 800001 910000
AR274513_09 900001 1010000
AR274513_10 1000001 1110000
AR274513_11 1100001 1210000
AR274513_12 1200001 1310000
AR274513_13 1300001 1410000
AR274513_14 1400001 1510000
AR274513_15 1500001 1610000
AR274513_16 1600001 1710000
AR274513_17 1700001 1810000
AR274513_18 1800001 1830121
Continuation (10 of 19) of AR274513 from base 900001 (AR274513 Sequence 1 from patent US

Query Match 17.7%; Score 292.4; DB 6; Length 110000;
Best Local Similarity 85.1%; Pred. No. 4.9e-38;
Matches 336; Conservative 1; Mismatches 57; Indels 1; Gaps 1;

8 ATCACAATAGSGGATCCACGAGCTTCTATTAGGTATCGTATTCGTCGAGAGGATATCC 67
41919 ATATCAATGGATCTCGAGCTTCT-TTAGGTATCGTATTCGTCGAGAGGATACCC 41977
68 AAAGGATTATCCAAAGCGCATCAATCAGCGATTCCTTAAAGTCGCTCAAAAACGA 127
41978 AAAGATTATCCAAAGCGCATCAATCAGCGATTCCTTAAAGTCGCTCAAAAACGA 42037
128 GAAAGTTTCTTAGCGGGTGTGCGAGAACAAAGAGCAAGCTAGTCACAAACGCGGTCG 187
42038 GAAAGTTTCTTAGCGGGTGTGCGAGAACAAAGAGCAAGCTAGTCACAAACGCGGTCG 42097
188 TGTACTTTGTGTGCTGCTAGCGGTAAGAGTGTATTGAAGCAACAACAAAGCGTTAAA 247
42098 TGTACTTTGTGTGCTGCTAGCGGTAAGAGTGTATTGAAGCAACAACAAAGCGTTAAA 42157
248 ATTGGCTGAGCAAAATTCATGGTCTGGGGCTTTTATCGTCGAGACATTTGTTTACAGGC 307
42158 ATTGGCTGAGCAAAATTCATGGTCTGGGGCTTTTATCGTCGAGACATTTGTTTACAGGC 42217
308 TGTGGAACGAGAACCAAGCAAAAATAGTTAGAAATCTTTGTTGAATTTAATAGATAAAAAAT 367
42218 TGTGGAACGAGAGTTCACAAAATAATCCATATCATTAATGAATTAATTTCAATTAAGT 42277
368 ATTGTACAGGGTAGAATTTGTTATTTTCTAGGATTT 402
42278 GATGAAAATTTAGGAATGAATTTTAAATTTAAATTT 42312

RESULT 4
AE006056
LOCUS
AE006056 11122 bp DNA linear BCT 16-JUN-2004

/translation="NNILIGNGGRHEHALAKVQSPMAEKVFPVAPGNAGTALBEKVE
NVAISATDVEKILVPAQAHIGHLITVGPPEALVIGVDPAFRAAGIKI GPPTKAAQOLE
GSKATKDFLARHPIPTAEYQNFTEVEPALAVREKGAPIVVKADGLAAGKGVIVAMT
LDEAAAKVEMSGNAFGAGSRVVEFELDGEERAFVWVDGKGNVEPMASQDKV
GENDTGLANTGNGAGSPAPVTPPEIHERIMQOYIPTVNGMAEGNIYTGFLYAGLMI
MNGAPKVFIEFCRFGDEPTQPMURLESIDLVELCIAACDGKLDQKKSQWCEQASLGI
VLAEGYFGDVRKGDDEITGIBSAVENQKVFVLAGVENKDGKLVNTNGRVRVVCVTALGDTV
YDAQQALALAEQVWKTRFYRRDIDIGYRAVAREKNA"
9760..11022
/gene="glyA"
/note="synonym: PM0225"
/760..11022
/gene="glyA"
/codon_start=1
/transl_table=11
/product="GlyA"
/protein_id="AAK02309.1"
/db_xref="GI:1270450"

gene

CDS

/translation="MLKRSNMIADYDVLWQAIQDENRROBEHIELIASENYASPRYM
EAQSQFTNKYAEYGPGRYVYGCCEYADIVQALAIQDRAKELFHADYVNVOPHSGSQAN
AAVYGALLQPHDTILGMSLAHGHGHLTHGSAVSFGKIYNAVOYGITAEGLIDYEDVRQ
KALECKPMIVAGFSAYSOVVDMAKREIADVEGAYLFVDMAHVAGLIVAGVYPSLP
HAHVVTTHHTKLGPRGLILSAKODELYKLOQSSVFPANQSGPLVHVIKAVCF
KEALPEYKVYQQQVVKAKAMVDVFKORGXNVVNGTENHLFLVDLVSHGLTGKAAD
AALGSANIVNKNVNDPQKPFVTSGRVGTSPSITRRGFKEABESBELAGMWCVDLDA
MKONEAQVIAQTEKVLAIKRLPVA"

ORIGIN

Query Match 17.4%; Score 288.4; DB 1; Length 11122;
Best Local Similarity 56.1%; Pred. No. 3.3e-37;
Matches 626; Conservative 4; Mismatches 465; Indels 21; Gaps 4;

QY	499	TTGTCATCGTAGTATTATATCTGCAATTCTCTTTTCATTAGAAAGTTATTATATAGAAAAT	558
DB	6767	TTTTTATTTTATGGCGTATTCGCGCTTTTTCGCTATTGAGCGAAATACCTTTATCGTTATT	6826
QY	559	TATTTAAATTTCTAGTATTGAGAGATACACTGAAAGTTATTGCTGTGTTGTTGTTG	618
DB	6827	TTTTCAAGATTAAATAGCTATTACCGTATTAGTGAAGTTTCTGATTATCTTTGTTGTC	6886
QY	619	TTTGTTGTTGTTGTTTTCAAAATATAGAAATTAAGAAATATTAGTTGGCGCTTTATTG	678
DB	6887	TTTCATTAATTTATTTTCTAAATATAGGGTATCACGCTTTTATCGCTTTATCTCTTT	6946
QY	679	CAATAAGTATTGTTGTTAATGATGACATTATGAGTATACCAATCTTGGATTGGACCTG	738
DB	6947	CTTGAGTAGCTGATTATTAACGTCGATTATGAGTTTATCAAAACTGGATTAAATGCGA	7006
QY	739	TTAATTACTCTGCTGCTTTAAAGAAATATGAGATAACAAATGCTGGCTTAACAATGA	798
DB	7007	CGAATTATCTACTAATGTTTAAAGAAATATGCGAAGTCAACCATGCGGTTTACATATGT	7066
QY	799	TAGATAAATTCATATATCATCTGTTATTGTTGTTTATTGAAAGTGTGCTGTTTAAAGTT	858
DB	7067	TAGCAAACTTGGCGTGGCATTATTGGGGGGCGCTAGA-----TATCTGATT	7117
QY	859	TAAGTTTCATAAAAAGAAAGTATATAAATCTTCTGGAATTTTGACATTATTTTTTATG	918
DB	7118	TATTTCCATATATCGGTTTATAGGCAAAAACCGATTGGGTGGCTGATATTTCGTTTATCT	7177
QY	919	CTGTGATGATGATGTTTGTGTTGCGAGGTATACAAACAAATCCATGAGCGTTTATTT	978
DB	7178	TGGTATGGGCTATATCTTTGCGGCTCTTTTATACCAA---TCAAGGTGAGGATCA	7234
QY	979	CACCTAACACTGTTTATCTCGATTAATAATCCAAATTTATCTGTTGGGTTATTTATAG	1038
DB	7235	CCTCTAATCCTGGATATCGAGATTAAAGCAACTTTTTCGCTTTGGTTACTTTATG	7294
QY	1039	GACGAATGTTCTTATGAGATATTTCTTTATCTAATATCTCTTTTATCATAAATCTA	1098
DB	7295	GTAATAACACTGCTTATGATCTCTTTAATTTAGTAATATGCTGTGTTATATCGAGATA	7354
QY	1099	AGCCTATGAATCGGCTCTCCGVAAATAATTMARGAATATAATTTTAAATTAATGGGGAA	1158

Db	7355	AACCTAATATGACACAGCCTCC---CATGCAAAAATAATTGCTCTTAATTATGGAGAA	7411
QY	1159	GTGGACCTCAAGTCATTTTAGTGCTTTTGGTTACGGGAGAAAACAATCTCTCTTTTATG	1218
Db	7412	GTTTGAGTGCAGCAATGTAGGACGTTTGGCTATGAAAGGCAACCATGCCATTTTGG	7471
QY	1219	ATAGCTTAAAAATAAATC-----AGGAGCTCTGTTGTGTAATACTTATTCAGGAGAA	1272
Db	7472	ATCAGTTAGTCAAAATGCAACCGCGAGAACGCTATTGAAAACAGCCTATTCTGCTGGT	7531
QY	1273	AGCTAACAGCAATTTCTTTTACCAATGTTTTTAAATGCAATTCCTTAYCCAAATGGAATAC	1332
Db	7532	TAGTAGCGGCTATTTTCATTGCTGCTTTTAAATGCAATTCCTCGTCCAAATGGCTAG	7591
QY	1333	AACAGATAGCTAAGAGATACGAATTTTAAATTTAGGAAAAGAGAGGCTTTTCAGA	1392
Db	7592	AACAAATTTGTAGTGCAGAACAACTTTGTCGTTTAGCGAAGAACGAGGCTATCAGA	7651
QY	1393	CATATTTTATTCAGCTCAAGCTAGGATGATATGATCATATGATCAATTTTATGAGGAGG	1452
Db	7652	CCTATTTTATTCAGCGCGCTGAAAACCAATGATGATCATGATATCATCGGGAATA	7711
QY	1453	CTTGATTTGATGATATTCGTTTCCAGATAATGAAGGGTATTCTTTTAAGAGATTTCAATGC	1512
Db	7712	CTGCGGTTGATCATTTTACTTTTCCCGAGTGATCGCTATAAACGTTCCAGAGGAATGC	7771
QY	1513	CTGATATAAATTTACTCTGCTGCTTTTAAAAATTAATTTAGATATGTTATCATTTTG	1572
Db	7772	ATGATCATGCCCTTCTCCCTTTTATTTGAACAAATAGATTGAGTGAAGGGAATCATTTTA	7831
QY	1573	TTGTTTATCATATAGAGGAGTCAATTTCCCTATG	1608
Db	7832	TCGTGTTATCAAGAGGATCTCAGCGACCTATG	7867

RESULT 5
AX024062

LOCUS AX024062 1572 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 5 from Patent FR2785293.
ACCESSION AX024062
VERSION AX024062.1 GI:10184374

KEYWORDS

SOURCE

ORGANISM

Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

Neisseriaceae; Neisseria.

REFERENCE

AUTHORS

Nassif, X., Tinsley, C., Aujaime, L., Perrin, A., Rokbi, B.,

Bouchardon, A. and Renaud, M.G.

Patent: FR 2785293-A 5 05-MAY-2000;

PASTEUR MERIEUX SERUMS VACC (PR)

FEATURES

Location/Qualifiers

1..1572

/organism="Neisseria meningitidis"

/mol_type="unassigned DNA"

/db_xref="taxon:487"

1..1572

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CAC08907.1"

/db_xref="GI:10184375"

CDS

translation="MKKSLFVLYSSLLTASEIAYRFVFGIETLPLPAKMAETPALTF
MIAALVLPARYKASSELLIAVFAFASIANNVHVAVYQSMVTGINVLMKLETVEGSA
GASMLDKLWLPALWGVLEWMLFCSLAKFHKRTHFSADILFALMLMIVFRSFDTKQEH
GLSPKPTSRKANKYFSGYFVGRVLPQLDFSLRIPAFKQFAPSKIQGSSVQNVLI
MSESAAHLKLFYGRFETSPFLTSLQADFKPIVKQSYAGFMTAVLSLPSFNAIPH
ANGLEQISGGDTNMFRLAKQEGYETFYSAQENEMAILNLIGKWIIDLHILQPTQLGY
NGDNMPDKLLPLDKLNLQGRHFIVLHQGRSHAPYSALLQPODKVFGELIVKDY
NTIHKTDQIOTVPEOLQKQDGNLWLFAYTSDHGOYVRQDIYNGTVOPODSYLVPLVL
YSSNKAVQQAANQAFAPCEIAPHQOLSTFLIHTLGYDMPVSGREGSVTGNLITDAG
SLNRDGRKAEVTFQ"

ORIGIN

Query Match 11.9%; Score 197.8; DB 6; Length 1572;
Best Local Similarity 51.6%; Pred. No. 2e-22;
Matches 533: Conservative 2; Mismatches 480; Indels 18.

[illegible]

Qy 1606 ATGGGGCATTATT 1618
1115 ACAGCGCATTGTT 1127
Db

RESULT 6

A96272	LOCUS	A96272	1575 bp	DNA	linear	PAT 07-SEP-2000
	DEFINITION	Sequence 305 from Patent WO9324578.				

ACCESSION A96272
VERSION A96272.1 GI:6780028

KEYWORDS	
SOURCE	
ORGANISM	
.	unidentified
	unidentified
	unclassified.

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Piazza, M., Scariato, V., Rappuoli, R., Grandi, G. and Masignani, V.
Neisseria antigens
Patent: WO 924578-A 305 20-MAY-1999;
PIZZA MARIAGRAZIA (IT); SCARIATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)

```

FEATURES
  source
    1..1575
      /organism="unidentified"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32644"

```

ORIGIN

Query Match	11.8;	Score	196.2;	DB	6;	Length	1575;
Best Local Similarity	51.5;	Pred.	No. 3.6e-24;				
Matches	532;	Conservative	2;	Mismatches	481;	Indels	18;
						Gaps	3;

[illegible]

[illegible]

QY 1426 TGCAATATGATCAATTTTATAGGAGGAGCTTGATGATGATATTCGTTTCCAGATAATG 1485
Db 935 TGGCAATTTTGAATCTTAATCGGTAAAGAAATGGATAGACCATTCGATTCAGCGCAGCAC 994
QY 1486 AAGGGTATCTTTTAAGAGATTCAATCCCTGATAATAAATTAATCTCTGCTTTTAAAAATA 1545
Db 995 TTGGCTACGGCAACGGCGACATATGCCGATGAGAGCTGCTGCGTTGTTTCGACAAA 1054
QY 1546 TTAATTTAGATATGTTATCATTTTGTGTTTATCATCATGAGGAGTATTCCT 1605
Db 1055 TCAATTTGAGCAGGCGCAGGCAATTTTATCGTTGTCACCAACGCGTTCGCAGCGCCCAT 1114
QY 1606 ATGGGCAATATT 1618
Db 1115 ACGGCGCATTTGT 1127

RESULT 12
AP005083 288108 bp DNA linear BCT 07-APR-2004
DEFINITION Vibrio parahaemolyticus DNA, chromosome 1, complete sequence,
11/11.
ACCESSION AP005083 BA000031
VERSION AP005083.1 GI:28807856
KEYWORDS Vibrio parahaemolyticus
SOURCE Vibrio parahaemolyticus
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
REFERENCE 1
AUTHORS Nasu,H., Iida,T., Sugahara,T., Yamaichi,Y., Park,K.S., Yokoyama,K.,
Makino,K., Shinagawa,H. and Honda,T.
TITLE A filamentous phage associated with recent pandemic Vibrio
parahaemolyticus O3:K6 strains
J. Clin. Microbiol. 38 (6), 2156-2161 (2000)
20235086
10834969
REFERENCE 2
AUTHORS Makino,K., Oshima,K., Kurokawa,K., Yokoyama,K., Uda,T.,
Tagomori,K., Iijima,Y., Najima,M., Nakano,M., Yamashita,A.,
Kubota,Y., Kimura,S., Yasunaga,T., Honda,T., Shinagawa,H.,
Hattori,M. and Iida,T.
TITLE Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V cholerae
Lancet 361 (9359), 743-749 (2003)
22508454
12620739
REFERENCE 3 (bases 1 to 288108)
AUTHORS Hattori,M., Yamashita,A., Oshima,K. and Shiba,T.
TITLE Direct Submission
Submitted (09-APR-2002) Masahira Hattori, Kitasato Institute for
Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara,
Kanagawa 228-8555, Japan
(E-mail:hattori@genome.ls.kitasato-u.ac.jp,
URL:http://genome.ls.kitasato-u.ac.jp/, Tel:81-42-778-8194,
Fax:81-42-778-8193)

COMMENT This work was done in collaboration with Koza Makino, Ken Kurokawa,
Katsushi Yokoyama, Takayuki Uda, Kenichi Tagomori, Masatomo Najima,
Masayuki Nakano, Yoshino Kubota, Shigenobu Kimura, Teruo Yasunaga,
Takeshi Honda, Hideo Shinagawa, Tetsuya Iida (Osaka University),
Yoshio Iijima (Kobe Institute of Health), and supported by the
Research for the Future Program of the Japan Society for the
Promotion of Science. This clone was isolated from a patient
presenting with acute gastroenteritis.

FEATURES
source Location/Qualifiers
1. .288108
/organism="Vibrio parahaemolyticus"
/mol_type="genomic DNA"
/strain="O3:K6"
/sub_strain="R1MD 2210633"
/db_xref="taxon:670"
/chromosome="1"
91. .555
gene

CDS
/gene="VP2831"
91. .555
/gene="VP2831"
/note="similar to GB:AAF95794.1 (AE004331) percent
identity 92 in 150 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="protein-transport protein SecB"
/protein_id="BAC61094.1"
/db_xref="GI:28807857"
/translation="MAEAPQEAQONFAIORIFLKVDVSEAPNSPVIQKWNPDVKL
DLDTQSRELGGVYEVRLRLVTVYKNEBETAFCEVQGGIFTAEQMEAGQLAHLCLGA
FCPNILFPYARETISLVLVVGKTFPOLNAPVNFDAFMNLIQQQOQGEAQA"
752. .1789
/gene="VP2832"
752. .1789
/gene="VP2832"
/note="similar to GB:AAF95792.1 (AE004331) percent
identity 86 in 343 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="glycerol-3-phosphate dehydrogenase (NAD+)"
/protein_id="BAC61095.1"
/db_xref="GI:28807858"
/translation="MTQANTNNAYGKDIAMTVIGAGSYGTSLSIARNGANVVLWGH
EPHMRLEADRANHAFLPGVDFPESLIIESDLKAVQASRDLLVVVPSHVFGIVLNS
CKPFLRDSRICWATKFLPETGRLLKQVADIIGENSLAVLSGPTFAKELAMGLPT
AISVASPAEFVADLQEKIHCSTKFRVYANNDFIGMOLGGAVKNVIAIGAGMSDGI
GALADLTITGLAESRLGALGAKPFTFGMAGLGLVLTCTDQNRNRRPGLALG
QGQVDVTAQEEIGQVVEGYRNTKEVMLSLRGKGVEMPIVDQIYVLYQGLDARLAAQD
LLARDKKAEGK"
1868. .2689
/gene="VP2833"
1868. .2689
/gene="VP2833"
/note="similar to GB:AAF95790.1 (AE004331) percent
identity 88 in 273 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="serine acetyltransferase"
/protein_id="BAC61096.1"
/db_xref="GI:28807859"
/translation="MKHCEKQKWNKIVSEAREMSEQEPMLASFYHATIIKHESLCAA
LSYLLANKLNTASMPAMAVREVEAFADPTITDFAACDIDCATVNRDPAVSMYSMPL
LYLKGYHALQGRVYVNWVKQRHALATYLNQISVACQVDIHPAARIGSGIMLDBAT
GIVIGETAVIDNDVSILOQDVTLLGGTGKCGDRHPKIRGVMIGAKILGNIEVGEGA
KIGCSVVLOAVPPHTTVAGVPAKIVGRPKTDKPSLMDQGFNGKSQSFHGDGI"
2755. .3861
/gene="VP2834"
2755. .3861
/gene="VP2834"
/note="similar to GB:AAF93508.1 (AE004122) percent
identity 65 in 362 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="NipD-related protein"
/protein_id="BAC61097.1"
/db_xref="GI:28807860"
/translation="MLRGAILLTCAISATYPLSSFAAQELKGSVSEINRQKALTS
QEKOLNEQSLKDQELGSKLERIKOTKADLAQADKNINKLESLSQEAQAOAE
HNKQLELEKSGIETLLKQSEKRTALASTQSNRKQTLNKIQSSINDKRYLAELORN
ETRLKAEIAKAKRNAVPMGLGKQKGLPWLKGSVLHNFQTRGTQGNVWKGVLISA
NYGQVKAIVPCTVVFVYGLVGLVLLDHGKGMTLYGNQALTKEGDKVTAGEV
IALAGDTGGDRPSPLYFEIRNSEAQNPKSWLKR"
complement (3938. .4894)
/gene="VP2835"
complement (3938. .4894)
/gene="VP2835"

/note="similar to GB:AAF93507.1 (AE004122) percent
identity 56 in 318 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC61098.1"
/db_xref="GI:28807861"
/translation="MNYDIFNGDADGIIALLQLRLADPIDSLITGVKRDIKLVEKVD
VQADELITVDISHEKMKAGLEQALAGAHVFYDHHKAGDI FOHGNLDAHIDLNNM
CTALIVDKLLEGRHTWAIYAAAGNDLIAKADVLADQAGLNNEQKQKELGFLINYN
GYGKVDLDFHFDLRYALVQYISPEVIEDKASTYIQQSAYQQDDMAAQVAPATH
ESDTLKLFEPLNPAASRSISGVYGNWLANQNPDSAHAVLTENADGTIVTSLRAPLNNK
QGA VAVCGQPTGGGREGAAAGINALRKEDYNFIDAVETYYA"
complement(5042. .5677)
/gene="VP2836"
complement(5042. .5677)
/gene="VP2836"
/note="similar to GB:AAF93506.1 (AE004122) percent
identity 74 in 209 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="transcriptional regulator, Tetr family"
/protein_id="BAC61099.1"
/db_xref="GI:28807862"
/translation="MKTRDKIVYAAELFNOHGERNITWNHIAHLEISPGNLYVHFR
NKQIVREIFALYSALLERFTPIQGSQESLTMKSYLDSIFITMKYRFFANLPEI
LSRDEQHEQTIDVQEKLANLAIMQEFVSLKLLDVNEQOLKSLVCTHLIACSWLA
YQSAMASKTITEQMVKQMLQMLNVVYKPVATEQGLEQLQLLEAVSTLQG"
complement(5699. .6073)
/gene="VP2837"
complement(5699. .6073)
/gene="VP2837"
/note="unknown"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC61100.1"
/db_xref="GI:28807863"
/translation="MNIYTAARFVILSYAILIAFLISLSSGELQWGNIAVLDMQOHAMA
YSGFAPFGCFVWRWRQWAVAMGILLFSVIEWLQGVYGRQTSWLDLVANSTGILT
LLCYRIWSLYRPLRVKLSYIR"
complement(6083. .7054)
/gene="VP2838"
complement(6083. .7054)
/gene="VP2838"
/note="similar to GB:AAF95796.1 (AE004331) percent
identity 83 in 321 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="lysyl-tRNA synthetase-related protein"
/protein_id="BAC61101.1"
/db_xref="GI:28807864"
/translation="MQNWQPTASIEQLRORATLILAIROFFAERQVMEVDTPAMSHA
TVDIHHTQTEFVGEGYADGSKLFMTSPFHMKRLAAGSGCIYQINAKFNBEN
GRVNPFTWLEWYRGFDHKLMDMDLLQLVLCGAAQRMVYQQAFTDVLGVCP
EGSMTELKAAASKIGLSDIAEPEDRDTLLQLLFSVGVENKIQDVPYAFYDFPASQA
ALAKINPDHRVADRFVYFKGIELANGPHELDNPKQLARPEQDNAKRIEMGLKPKP
IDVHLISALBAGLPDCAGVALGIDRLIMLALGCDHIDQVTAFFPPIA"
7158. .7280
/gene="VP2839"
7158. .7280
/gene="VP2839"
/note="unknown"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC61102.1"
/db_xref="GI:28807865"

gene
CDS

gene
CDS

gene
CDS

gene
CDS

Query Match 8.1%; Score 134.4; DB 1; Length 288108;
Best Local Similarity 61.0%; Pred. No. 1.4e-12;
Matches 233; Conservative 1; Mismatches 147; Indels 1; Gaps 1;
QY 12 CAATAGSGGATCCACGAGCTTCCTATTAGTATCGTATTCGCTGCGAGGATATCCAAAG 71
DB 82880 CTAATGGGATCCACGCGATCTA-TCGGTATCGTGTAGCAGCGGTCGCTACCCAGCA 82938
QY 72 GATTATCGCAAGGCGATGAATTCAGCGGATTCCTTAAAGTCGGTCAAAACGAGAAA 131
DB 82939 GCGTACACAAAGGTGACGTGATCTCTGCGCTGCCTCAAAGTTGAAATTTGAAGCGGAGAA 82998
QY 132 GTTTTCTTAGCGGTCGTCAGAACAAAGGCAAGCTAGTCACAAACGCGCGTCTGTGTA 191
DB 82999 GTCTTCACGCGGTCACAGCAATCAAGATGGGACATCGTGACAATGGTGGTCGGTA 83058
QY 192 CTTTGTGTGACTGCGTTAGCGGAAAGTGTATTGAGACACAAACAAACGCTTAAATTTG 251
DB 83059 TTGTGTGCAACTGCATTGGGCAACGCGTATCAGAAGCACAAACGCGCATACGAGCTA 83118
QY 252 GCTGACCAAAATTCATGGTCTCGGCGCTTTTATTCGTCGAGACATTGGTTACAGGGCTGTG 311
DB 83119 GCAAAACAGATCAGTTGGGATGCGCATGTTCCATCGTAATGACATCGGCTACCGTGTATC 83178
QY 312 GAACGAGAACAAAGCAAAATAGTTAGAAAATCTTGTGAAATTTAATTAAGATAAAAAATATTG 371
DB 83179 GCTCGGAGACAGAAAATAAATAACAACTTCAGTTGGTGAAGACAAAAGCGCATTC 83238
QY 372 TACAGGTTAGAATTTGTTATTTTC 393
DB 83239 AGCGCTTTTATTTGCTCTGTC 83260
RESULT 13
AE017153/c 300521 bp DNA linear BCT 15-AUG-2003
LOCUS
DEFINITION Haemophilus ducreyi strain 35000HP section 3 of 6 of the complete genome.
ACCESSION AE017153 AE017143
VERSION AE017153.1 GI:33148155
KEYWORDS
SOURCE Haemophilus ducreyi 35000HP
ORGANISM Haemophilus ducreyi 35000HP
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.
REFERENCE 1 (bases 1 to 300521)
AUTHORS Munson,R.S. Jr., Ray,W.C., Mahairas,G., Sabo,P., Munger,R., Johnson,L., Nguyen,D., Wang,J., Forst,C. and Hood,L.
TITLE The Complete Genome Sequence of Haemophilus ducreyi
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 300521)
AUTHORS Munson,R.S. Jr., Ray,W.C., Mahairas,G., Sabo,P., Munger,R., Johnson,L., Nguyen,D., Wang,J., Forst,C. and Hood,L.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2003) Pediatrics, Columbus Children's Research Institute and The Ohio State University, 700 Children's Drive, Columbus, OH 43205, USA
FEATURES
source 1. .300521
/organism="Haemophilus ducreyi 35000HP"
/mol_type="genomic DNA"
/strain="35000HP"
/db_xref="taxon:233412"
90. .614
/locus_tag="HD0771"
90. .614
/locus_tag="HD0771"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP95676.1"
/db_xref="GI:33148156"

/translation="MOKVKLPITIDPFKDAQRRVDYGVYFSSKQLVRLGESVCNVIND
AQVTLISYVDPQRLSVLKGSAQVEVELCCQRCNSFTQIDCSFLFSPVRNLDQAGDL
PIYPIEYVNAFGEVNLDMIDEFTLALPLVPMHSEAHCEVSVSSEQVFGELPABLAK
KNPFAVLNKKN"
633. .803

gene

/gene="rpmF"
/locus_tag="HD0772"

CDS

633. .803
/gene="rpmF"
/locus_tag="HD0772"

CDS

/codon_start=1
/transl_table=11

/product="50S ribosomal protein L32"

/protein_id="AAP95677.1"

/db_xref="GI:33148157"

/translation="NAVQNKKSRRRDMRRSHDALTTAAVSVDKTTGTGTHLRHHVTA
DCYVGRKVIK"
825. .1856

gene

/gene="plsx"
/locus_tag="HD0773"

CDS

825. .1856
/gene="plsx"
/locus_tag="HD0773"

CDS

/codon_start=1
/transl_table=11

/product="fatty acid/phospholipid synthesis protein"

/protein_id="AAP95678.1"

/db_xref="GI:33148158"

/translation="NRRLTALDVMGDFGPRVTIPALSALAAQPMLSFVLPGDOYH
AIPHLSPAEQQRIRFVHTSRIVDHPVVRQALRQSGSSRWLAIEAVAGEAQAC
VSGGNTGVMLGAKFVEPLVIGDRPALASLIIPALNGQASVMDIGANVEADSKILLQ
FAEMGNIFAQSLAQRLALLNTGTEVKGRIYAIQQAHVLYKNRADLHYGFIESD
KLMLHLADVICDQFCGNIALKAVEGASHFLAKQPKRSFFVDWLSWNGRLLYHR
VYRALQVRNPERHNGTLLGLSSVVVKSHSGAGVNAYFYAINHAIEQIQNRIPVKISA
GLQKLHQNL"
1868. .2818

gene

/gene="fabH"
/locus_tag="HD0774"

CDS

1868. .2818
/gene="fabH"
/locus_tag="HD0774"

CDS

/codon_start=1
/transl_table=11

/product="3-oxoacyl- (acyl-carrier-protein) synthase III"

/protein_id="AAP95679.1"

/db_xref="GI:33148159"

/translation="WYSKILATGSLYLPQVRSNADLEKNVDTTDEWIFTRSCWKERRI
AADETVACTGAAAHKACEMAIDVNEIELIVAITTTHSHAYPSAAQIQGMLAIED
AIAFDVAACTGFIALSVADQFIRAGKYKATPVVGADVNSRVLDEADRGTVVLFQDG
AGAVTIQASEQAGILSTHLSSADTDNMLALPIQERGNASGFIQWGNATPKLAVSR
LANVVEETQANNLOKADLDMLVPHQANIRIIAATAKLEMEMSQVLTVEKYGNNSA
ATVPVALDEAVDRGRIQRQGLLLLEAFGGGWTWGSGLVRF"
2828. .3577

gene

/locus_tag="HD0775"

CDS

2828. .3577
/locus_tag="HD0775"

CDS

/codon_start=1
/transl_table=11

/product="probable methyl transferase"

/protein_id="AAP95680.1"

/db_xref="GI:33148160"

/translation="MSEKTFADQKRKTVEQAFTEDGRLRKVRFSVLRTGRLSDQFQ
RWMNDHMAKFLGDYQKSPDFVEIFGNANPVVLEIFGCMGRSLGVEMAQNPDRNVIG
IEVHTPGVACIAYAEKQVDNLRVICHDATEILRDADVADGSLAGLQLYFPDPWQSK
HKRRIRIQTVEFWRLTLTKTAGGFTHPATDQWQYAEHMLAVLRQSPALVNTSPTGDP
IDRPAPRLTKFEAGQRGLGHAVWDLIYFKK"
3658. .3999

gene

/locus_tag="HD0776"

CDS

3658. .3999
/locus_tag="HD0776"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0777"

CDS

3658. .3999
/locus_tag="HD0777"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0778"

CDS

3658. .3999
/locus_tag="HD0778"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0779"

CDS

3658. .3999
/locus_tag="HD0779"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0780"

CDS

3658. .3999
/locus_tag="HD0780"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0781"

CDS

3658. .3999
/locus_tag="HD0781"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0782"

CDS

3658. .3999
/locus_tag="HD0782"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0783"

CDS

3658. .3999
/locus_tag="HD0783"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0784"

CDS

3658. .3999
/locus_tag="HD0784"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0785"

CDS

3658. .3999
/locus_tag="HD0785"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0786"

CDS

3658. .3999
/locus_tag="HD0786"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0787"

CDS

3658. .3999
/locus_tag="HD0787"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0788"

CDS

3658. .3999
/locus_tag="HD0788"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0789"

CDS

3658. .3999
/locus_tag="HD0789"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0790"

CDS

3658. .3999
/locus_tag="HD0790"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0791"

CDS

3658. .3999
/locus_tag="HD0791"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0792"

CDS

3658. .3999
/locus_tag="HD0792"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATAAAYSVMALVPIQPIKALTTEQERKIRWOLRTVSAREKIVFVPLILL
VALLPDAAPLLGMCFGNLMRASGVVERLNATQALINIVTIFLGSVAGKLAVDK
FLQPTQLGLLGIIVAGTAGSVIMAKIMNAFSGKINPLIGSAGVSAVPMARVVS
NKIGLEADKQNFLLMHAMGNVAVGVISSAIAAGIMLKYSTMI"

gene

/locus_tag="HD0793"

CDS

3658. .3999
/locus_tag="HD0793"

CDS

/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"

/translation="MESIILAFQGMGIMHITLQGVIMIVSLILWLAIARPEPPLL
LPFGGLLSNIPEAGLMTALDHLHLYAHTQLTILAEKNSSSDVDHTTIKQAIATA
PSPVQNELEVIASDGLYNAGIALFYKVAIGVAPLIIFMGVGMATDFGLLANPRT
LLGAAQFGIFTTVLGLGNLWGIIDFSPQAAAIGIIGDAGDTATYILASKLAPE
LIGATA

Qy	649	TTTCAAGAAATATTAGTTGGCGCTTTATTTGCAATAAGTATTGTTGTTAATAATGTCACATT	708
Ds	170	CAACGGCGTTTGTGATTGGCGGTGTTTTCGCGTTTCAGCATTTATGGCAACAATGTCGATT	229
Qy	709	ATCAGTATACCAATCTTGATTTGGACCTGTTAAATTACCTCACCTTGCATTTAAAGAATAA	768
Ds	230	ACCGGTTTATCAAGCTGGATAACGGGCATTTAAATTATTTGGCTGATGCTGAAAGAGATTA	289
Qy	769	ATGAGATAACAAATGCTGGCTTAAACAATGATAGATAAAATTCATATATCCATTCTTATTTG	828
Ds	290	CCGAGTTGGCGGCGACGGGCGTCGATGTTGATTAAGTTTGGCTGCGCTGCGGTTGGG	349
Qy	829	GTTTATTGGAAGTTGCTGTGTTTTTTAAAGTTTAAAGTTTTCATAAAAAAGAAAGTATATAAAC	888
Ds	350	CGCTGTTGGAAGTCATGTTGTTTTCAGCGCTTCCCAAGTTCCGCGCTAAGACGCAT----	405
Qy	889	TTTCTTGGATTGTTGATTTATTTTTTATGCTGTGATGATGATGTTTTTGTTCGAGCGT	948
Ds	406	-----TTTTCTGCGGATTAAGTCTGTTGCTTCTCTAATGCTGATGATTTTCGTCGCTT	457
Qy	949	ATACAAACAAATCCCATGAGCGTTTTTATTTTCACTTAACACTGTTTATTCGATTTAAAT	1008
Ds	458	CGTTCGACACGAACACGCGTATTTTCGCCCAAACGGACATACGCCGCGATCAAG	517
Qy	1009	CCAATTATTATCGTTGGGTATTTTATAGGAGCAATTTGTTCTTATGAGATATTTTCTT	1068
Ds	518	CCAATTATTACAGTTCGGTTATTTTTCGCGACGCGTGTTCGCGTATCATGTTCTTGATT	577
Qy	1069	TATCTAATATCTCTTTATCATAAATCTAAGCTATGAATCGGGCTCTCCGVAATAAT	1128
Ds	578	TAAGCAAGATTCCTGTGTTTCAAAACAGCTCTCCAAAGCAGATTCGGGCAAGCGATATTC	637
Qy	1129	MARGAATAATTTTAAATTAATGCGGGAAGTCGACCTCAAGTCATTTAGTGTCTTTG	1188
Ds	638	AAATATGTCCTGA---TTATGGCGGAAGCAAGCGCGGCGCATTTGAAATTTGTTG	694
Qy	1189	GTTACGGGAGAAAAACATCTCTCTTTTT---AGATAGCTTAAATATAAATCAGGAGCTC	1245
Ds	695	GCTACGGGCGGAACTTCGCGGTTTTCGACCCAGCTTCGCAAGCGATTTTAAAGCCGA	754
Qy	1246	TTGTTGGTAAACTTATTCAGGAGGAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA	1305
Ds	755	TTGTGAAAAAAGTTTATTCGAGGCTTTATGACGCGAGTATCCCTGCCAGTTCTCTTTA	814
Qy	1306	ATGCAATCTCTTAYCCAAATGGAATCAACAGATAGCTAAAGAGATA	1353
Ds	815	ACGTCATACCGCATGCAACGGCTTGGAAACAAATCAGCGCGCGGATA	862
RESULT	15		
LOCUS	NMA22491/c		
DEFINITION	Neisseria meningitidis serogroup A strain 22491 complete genome; segment 2/7.	DNA	linear BCT 02-SEP-2002
ACCESSION	AL162753		
VERSION	AL162753.2		
KEYWORDS	GI:7379120		
SOURCE	Neisseria meningitidis 22491		
ORGANISM	Neisseria meningitidis 22491		
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.		
AUTHORS	1 (bases 1 to 349061)		
	Parkhill, J., Achtman, M., James, K.D., Bentley, S.D., Churcher, C.,		
	Klee, S.R., Morelli, G., Basham, D., Brown, D., Chillingworth, T.,		
	Davies, R.M., Davis, P., Devlin, K., Feltwell, T., Hamlin, N.,		
	Holroyd, S., Jagsels, K., Leather, S., Moule, S., Mungall, K.,		
	Quail, M.A., Rajandream, M.A., Rutherford, K.M., Simmonds, M.,		
	Skelton, J., Whitehead, S., Spratt, B.G. and Barrall, B.G.		
TITLE	Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491		
JOURNAL	Nature 404 (6777), 502-506 (2000)		
MEDLINE	20222556		

```

PUBMED      10761919
REFERENCE    2 (bases 1 to 349061)
AUTHORS      Parkhill,J.
TITLE        Direct Submission
JOURNAL      Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
              sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
              Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT      Notes:
              Details of N. meningitidis sequencing at the Sanger Centre are
              available on the World Wide Web.
              (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES     Location/Qualifiers
             1..349061
               /organism="Neisseria meningitidis Z2491"
               /mol_type="genomic DNA"
               /strain="Z2491"
               /db_xref="taxon:122587"
               /note="serogroup: A"
               /complement(24..206)
               /notes="ATR repeat; hmms hit to HMM ATR (1 - 183), score:
               310.39"
               /label=ATR
             209..212
               /genes="NMA0368"
             220..681
               /genes="NMA0368"
             220..681
               /genes="NMA0368"
             /notes="NMA0368"
             153 aa: contains four probable integral membrane domains"
             /codon_start=1
             /transl_table=11
             /product="putative integral membrane protein"
             /protein_id="CAB83669.1"
             /db_xref="GI:7379121"
             /db_xref="TrEMBL:Q9JWH7"
             /translation="MQEQRKSFPIVMLVSVLWIASLSNVAFYLGNGHSGMEGLTV
             LILGSPASIDIRYCAVYANYWLAIVLLARKKVVPVHAAPFWGLALVAFSVKAVYV
             DRAGNTSDIVRYGALFYLAFAVASIGTFAGKNKKAASAADGTRNDV"
             671..1492
               /genes="hemK"
             671..1492
               /genes="hemK"
             /notes="hemK"
             /notes="NMA0369, hemK, HemK protein, len: 273 aa; similar
             to e.g. HEMK ECOLI P37186 HEMK protein (277 aa), fasta
             scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
             P500092 N-6 Adenine-specific DNA methylases signature"
             /codon_start=1
             /transl_table=11
             /product="HemK protein"
             /protein_id="CAB83670.1"
             /db_xref="GI:7379122"
             /db_xref="GOA:Q9JWH6"
             /db_xref="TrEMBL:Q9JWH6"
             /translation="WTFDKWLGSLPKNEARMLLQYVSEYTRVQLLTRGGEMPDEV
             RQRADRLAQRNGEPAVILGARFYGRRFTVNPVSLIPRPETHLVEAVLARLPEN
             GRVMDLGTSGAVAVTVALEPDAFVRASDISPPALETKRNAADLGARVEFAYGSWF
             DTDMPSEGDWIIIVSNPPYIENGDKHLSQGLRFEPQIALTFDSGLSCIRTLAQCAP
             DRLAEGGFLLHEHPDQGAARVGLAENGFSGVETLPDLAGLDRVTLGKYMKHLK"
             831..840
               /genes="hemK"
             /notes="Core DNA uptake sequence: gcgcgtctgaa"
             /label=DUS
             1162..1171
               /genes="hemK"
             /notes="Core DNA uptake sequence: gcgcgtctgaa"
             /label=DUS
             1187..1207
               /genes="hemK"
             /notes="P500092 N-6 Adenine-specific DNA methylases
             signature"
             /complement(1279..1288)
             /notes="Core DNA uptake sequence: gcgcgtctgaa"
             /label=DUS

```

```

RBS
gene
CDS

1565..1568
1578..2966
/genes="NMA0370"
/genes="NMA0370"
/notes="NMA0370, probable integral membrane protein, len:
462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
P44640 hypothetical protein HI0325 (450 aa), fasta scores;
E(): 0, 49.8% identity in 464 aa overlap"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB83671.1"
/db_xref="GI:7379123"
/db_xref="GOA:Q9JWH5"
/db_xref="TrEMBL:Q9JWH5"
/translation="NNAVVAVIVMLVLSRVHVVLSTLIGAFVGVGAGMPLONIA
DAAGOVSOAGIIPVFNKLEGCAKIALSYAMLGAFAMAITHSLGPQOLAGAVVRKLNK
GMPDSVRSRGEGVVKLLSIIIVGMMSQNIPIPIAFIPNLVPLLLVFNRLKIDR
RLIACVITFGLVTYTFUPYGFGAIFLINEILGNHISAAPQLDKRNINWMAAIPAL
GNLAGLLAFVHYRPRLYQSNNDTAGNADANRPQPSAYRSLAAVAIAVCFATQL
MYEDSLVLGMLGPAVFWMLGVINRDKANDVFEGIKMMAMVGFIMAAQGFAGAVNNA
TGHIOPLVSSMAIIRGNSKMAALMLVVLVLTWMTGIGSSFTLPIIAAIYVPLCVGL
GSPPLATVAIVGTAGALGDAGSPASDSTLIGTMTGLNADQGHDIRDSVPTFIHYNIP
LIIAGWIAAMVL"
2967..3191
/genes="alyX"
/genes="alyX"
/genes="alyX"
/notes="NMA0371, slyX, SLYX protein homolog, len: 74 aa;
similar to SLYX_HAEIN P44759 SLYX protein homolog (73 aa),
fasta scores; E(): 0.0017, 33.8% identity in 74 aa
overlap, and SLYX_ECOLI P30857 SLYX protein (72 aa), fasta
scores; E(): 0.24, 32.4% identity in 68 aa overlap"
/codon_start=1
/transl_table=11
/product="SLYX protein homolog"
/protein_id="CAB83672.1"
/db_xref="GI:7379124"
/db_xref="TrEMBL:Q9JWH4"
/translation="MDAVQEPHRTLEIQSALQEDVIAGLNAMVAELRQTLDLQQA
QLRLIYQKQDRNPDAQSPYSRLRDEIPHY"
/complement(3271..3423)
/genes="NMA0372"
/genes="NMA0372"
/notes="NMA0372, unknown, questionable CDS, len: 50 aa"
/codon_start=1
/transl_table=11
/product="very hypothetical protein NMA0372"
/protein_id="CAB83673.1"
/db_xref="GI:7379125"
/db_xref="TrEMBL:Q9JWH3"
/translation="NFGRLSPLGSGDFFLAVVEYVPGFICLOGKTAYFRSGGKRPHRI
KGILSD"
3562..3571
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
/complement(3586..4356)
/genes="thiF"
/complement(3586..4356)
/genes="thiF"
/notes="NMA0373, thiF, probable ThiF protein, len: 256 aa;
similar to e.g. THIIF_ECOLI P30138 THIIF protein (251 aa),
fasta scores; E(): 0, 43.1% identity in 246 aa overlap,
and MOEBP_ECOLI P12282 molybdopterin biosynthesis MOEBP
protein. (249 aa), fasta scores; E(): 0, 43.9% identity in
244 aa overlap (note that N.m. does not have orthologs of
any other molybdopterin biosynthesis proteins). Contains
pfam match to entry PF00899 ThiF_family, ThiF family"
/codon_start=1
/transl_table=11
/product="ThiF protein"

```

misc_feature

gene

CDS

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2004, 15:49:34 ; Search time 151 Seconds
(without alignments)
7795.142 Million cell updates/sec

Title: US-10-698-235-1
Perfect score: 1656
Sequence: 1 agtgaatcaccaatagggg.....gggtgttggaaraataaacg 1656

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/6CTUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	292.4	17.7	1830121	4	US-09-557-884-1
2	292.4	17.7	1830121	4	US-09-643-990A-1
3	292.4	17.7	1830121	4	US-10-329-960-1
4	95.4	5.8	1335	4	US-09-328-352-2917
5	88.2	5.3	1377	4	US-09-252-991A-10413
6	88.2	5.3	1383	4	US-09-252-991A-10306
7	88.2	5.3	1410	4	US-09-252-991A-10749
8	74	4.5	1141	4	US-09-806-708B-22
9	73.6	4.4	1299	4	US-09-543-681A-3189
10	73	4.4	10467	4	US-10-204-708-2
11	71.6	4.3	19513	4	US-10-204-708-39
12	70	4.2	5152	4	US-10-204-708-74
13	69.4	4.2	19124	2	US-08-487-826B-13
14	67.8	4.1	741	4	US-09-489-039A-1
15	67.6	4.1	8961	4	US-10-204-708-80
16	66.6	4.0	1055	4	US-09-806-708B-23
17	66	4.0	6113	4	US-10-204-708-14
18	65.8	4.0	6306	4	US-10-204-708-50
19	64.6	3.9	658	3	US-08-998-416-595
20	64.4	3.9	6866	4	US-10-204-708-20
21	63.8	3.9	1141	4	US-09-806-708B-22
22	63.6	3.8	5501	4	US-10-204-708-38
23	63.6	3.8	6070	4	US-10-204-708-10
24	63	3.8	9347	4	US-10-204-708-35
25	62.2	3.8	6669	4	US-10-204-708-5
26	62	3.7	6583	4	US-10-204-708-26
27	62	3.7	11049	4	US-10-204-708-22

28	61.8	3.7	636	3	US-08-998-416-1137	Sequence 1137, Ap
29	61.4	3.7	6669	4	US-10-204-708-6	Sequence 6, Appl
30	61	3.7	11049	4	US-10-204-708-21	Sequence 21, Appl
31	60.6	3.7	5455	4	US-10-204-708-33	Sequence 33, Appl
32	60.4	3.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
33	60.2	3.6	11049	4	US-10-204-708-23	Sequence 23, Appl
34	60	3.6	5852	1	US-07-867-108-2	Sequence 2, Appl
35	59.8	3.6	6866	4	US-10-204-708-19	Sequence 19, Appl
36	59.4	3.6	3111	4	US-09-252-991A-10504	Sequence 10504, A
37	58.8	3.6	6156	4	US-10-204-708-60	Sequence 60, Appl
38	58.6	3.5	5666	4	US-10-204-708-30	Sequence 30, Appl
39	58.4	3.5	5562	4	US-10-204-708-63	Sequence 63, Appl
40	58.4	3.5	6040	4	US-10-204-708-69	Sequence 69, Appl
41	58	3.5	8093	4	US-10-204-708-32	Sequence 32, Appl
42	57.8	3.5	6317	4	US-10-204-708-11	Sequence 11, Appl
43	57.6	3.5	10619	4	US-10-204-708-3	Sequence 3, Appl
44	57.6	3.5	19513	4	US-10-204-708-40	Sequence 40, Appl
45	57.2	3.5	5610	4	US-10-204-708-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 17.7%; Score 292.4; DB 4; Length 1830121;
Best Local Similarity 85.1%; Pred. No. 8e-54;
Matches 336; Conservative 1; Mismatches 57; Indels 1; Gaps 1;

QY 8 ATCAANTAGSGGATCCACAGCTTTCTATTAGGTATCGTATTGGCTGCAGAGGATATCC 67

```
Db 941919 ATATCAAAATGGGATCTCGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGATACCC 941977
Qy 68 AAAGGATTATCGCAAGCGGATGAATCAGCGGATTCGCTAAAGTGGGTCAAAAACGA 127
Db 941978 AAAAGATTATCGCAAGCGGATGAATCAGCGGATTCGCTAAAGTGGGTCAAAAACGA 942037
Qy 128 GAAAGTTTCTTAGCGGGTGTGCGAACAAGAGGCAAGCTAGTCAAAACGCGGTGCG 187
Db 942038 GAAAGTTTCTTAGCGGGTGTGCGAACAAGAGGCAAGCTAGTCAAAACGCGGTGCG 942097
Qy 188 TGTACTTTGTGCTACTCGTTAGCGAAAGTGTATTTGAAGCACACAAAAGCGTTTAAA 247
Db 942098 TGTACTTTGTGCTACTCGTTAGCGAAAGTGTATTTGAAGCACACAAAAGCGTTTAAA 942157
Qy 248 ATTGGCTGAGCAAAATCAATGGTCTGGCGGTTTATCGTCGAGACATTTGGTTACAGGC 307
Db 942158 ATTGGCTGAGCAAAATCAATGGTCTGGCGGTTTATCGTCGAGACATTTGGTTACAGGC 942217
Qy 308 TGTGGAACGAGACAAGCAAAATAGTTAGAAATCTTGTGAAATTTAATTAGATAAAAAAT 367
Db 942218 TGTGGAACGAGAGTGTGCAAAATATCCATATCATATGATGAATTAATTTAAGT 942277
Qy 368 ATTGTACAGGAGTAGAATTTGTTTCTCCTAGGATTT 402
Db 942278 GATGAAAATTTAGGAATGAAATTTTAAATTAATTT 942312
```

RESULT 2

```
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
```

```
Query Match 17.7%; Score 292.4; DB 4; Length 1830121;
Best Local Similarity 85.1%; Pred. No. 8e-54;
Matches 336; Conservative 1; Mismatches 57; Indels 1; Gaps 1;
Qy 8 ATCAAAATAGSGGATCCAGAGCTTCTATTAGTATCGTATTGGCTGCAGAGGATATCC 67
Db 941919 ATATCAAAATGGGATCTCGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGATACCC 941977
Qy 68 AAAGGATTATCGCAAGCGGATGAATCAGCGGATTCGCTAAAGTGGGTCAAAAACGA 127
Db 941978 AAAAGATTATCGCAAGCGGATGAATCAGCGGATTCGCTAAAGTGGGTCAAAAACGA 942037
Qy 128 GAAAGTTTCTTAGCGGGTGTGCGAACAAGAGGCAAGCTAGTCAAAACGCGGTGCG 187
Db 942038 GAAAGTTTCTTAGCGGGTGTGCGAACAAGAGGCAAGCTAGTCAAAACGCGGTGCG 942097
Qy 188 TGTACTTTGTGCTACTCGTTAGCGAAAGTGTATTTGAAGCACACAAAAGCGTTTAAA 247
Db 942098 TGTACTTTGTGCTACTCGTTAGCGAAAGTGTATTTGAAGCACACAAAAGCGTTTAAA 942157
Qy 248 ATTGGCTGAGCAAAATCAATGGTCTGGCGGTTTATCGTCGAGACATTTGGTTACAGGC 307
Db 942158 ATTGGCTGAGCAAAATCAATGGTCTGGCGGTTTATCGTCGAGACATTTGGTTACAGGC 942217
Qy 308 TGTGGAACGAGACAAGCAAAATAGTTAGAAATCTTGTGAAATTTAATTAGATAAAAAAT 367
Db 942218 TGTGGAACGAGAGTGTGCAAAATATCCATATCATATGATGAATTAATTTAAGT 942277
Qy 368 ATTGTACAGGAGTAGAATTTGTTTCTCCTAGGATTT 402
Db 942278 GATGAAAATTTAGGAATGAAATTTTAAATTAATTT 942312
```

RESULT 3

```
US-10-329-960-1
; Sequence 1, Application US/10329960
; Patent No. 6742927
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,960
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ IDS NOS: 1
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
```

```

; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
```

```

; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a, t, g or c
```

FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature

Query Match 17.7%; Score 292.4; DB 4; Length 1830121;
Best Local Similarity 85.1%; Pred. No. 8e-54;
Matches 336; Conservative 1; Mismatches 57; Indels 1; Gaps 1;
Qy 8 ATCAATAGSGGATCCAGGCTTCTATTAGGTATCGTATTGGCTGCAGGGGATATCC 67
Db 941919 ATATCAATAGSGGATCCAGGCTTCTATTAGGTATCGTATTGGCTGCAGGGGATATCC 941977
Qy 68 AAAGGATTATCGCAAGGCGATGAATCAGCGGATTGCTTAAAGTCCGCTCAAAAACGA 127
Db 941978 AAAAGATTATCGCAAGGCGATGAATCAGCGGATTGCTTAAAGTCCGCTCAAAAACGA 942037
Qy 128 GAAAGTTTCTTAGCGGGTGTGCGAGAACGAAGCAAGCTAGTCAAAACGCGGTGCG 187
Db 942038 GAAAGTTTCTTAGCGGGTGTGCGAGAACGAAGCAAGCTAGTCAAAACGCGGTGCG 942097
Qy 188 TGTAATTTGTGACTCGGTAGCGAAGTGTATTTGAAGCACAAACAAAGCGTTAAA 247
Db 942098 TGTAATTTGTGACTCGGTAGCGAAGTGTATTTGAAGCACAAACAAAGCGTTAAA 942157
Qy 248 ATTGGCTGAGCAAAATCAATGGTCTGGCGGTTTTTATCGTCGAGACATTTGTTACAGGC 307
Db 942158 ATTGGCTGAGCAAAATCAATGGTCTGGCGGTTTTTATCGTCGAGACATTTGTTACAGGC 942217
Qy 308 TGTGGAACGAGAACCAAGCAAAATAGTTAGAAATCTTGTGAATTTAATTAGATAAAAAAT 367
Db 942218 TGTGGAACGAGATTGCAAAATATCCATATCATATGATGAATTAATTTTCAATTTAAGT 942277
Qy 368 ATTGTACAGGGTACAATTTGATTTTCTTAGGATTT 402
Db 942278 GATGAAAATTTAGGAATGAATTTTAAATTAATTT 942312

RESULT 4
US-09-328-352-2917
; Sequence 2917, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2917
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2917

Query Match 5.8%; Score 95.4; DB 4; Length 1335;
Best Local Similarity 57.6%; Pred. No. 8.6e-12;
Matches 171; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
Qy 35 ATTAGGTATCGTATTGGCTGCAGAGGGATATCCAAAGGATTATCGCAAGCGGATGAAT 94
Db 1038 AATCGGTATTGTACTTTGCAGCAGAGAGGTTACCCAGACAGCGTGCCTAAAGGTACGCGTAT 1097
Qy 95 CAGCGGATTCCTTAAAGTGCCTCAAAAACGAGAAAGTTTCTTAGCGGGTGTGCGAGA 154
Db 1098 CTCAGGTATCGGTCAATTCACAGAGATATCAAAATCTTCCATGCGAGGCATCTCTACTCG 1157
Qy 155 ACAAGAAGCAAGCTAGTCACAAACGCGGTCTGTGTACTTTGTGTGACTGCGTTAGGCGA 214
Db 1158 TGAAGATGGCATATCGTTACTTCTGCGGACGTGTACTTTGCTGACTGCTCTAGGTGA 1217
Qy 215 AAGTGTATTGAGACACAAACAAAGCGTTAAATTTGGCTGAGCAAAATCAATGGTCTGG 274
Db 1218 TAGCGTTCTTGAAGCACAAATCAATGCTTTAGAAGTATGTGGCAAGTAACCTTCACAGG 1277
Qy 275 GCGTTTTTATCGTCGAGACATTTGGTTACAGGCTGTGGACGAGACAGCAAGCAAAATA 331
Db 1278 TATGCAATACCGCAGTGACATTTGGTTACCGTGTATTGCTGCGTGAAGAAAGCTGAATA 1334

RESULT 5
US-09-252-991A-10413
; Sequence 10413, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10413
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10413

Query Match 5.3%; Score 88.2; DB 4; Length 1377;
Best Local Similarity 56.8%; Pred. No. 3.2e-10;
Matches 162; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
Qy 37 TAGGTATCGTATTGGCTGCAGAGGGATATCCAAAGGATTATCGCAAGCGGATGAATCA 96
Db 1079 TGGGCGTGTACTGGCCCGCGCGGTATCCCGGCGACTACGCCAAAGGGCGAGGTCTATCG 1138
Qy 97 GCGGATTGCTTAAAGTGCCTCAAAAACGAGAAAGTTTCTTAGCGGGTGTGCGAGAAC 156
Db 1139 AAGGCTGCGCGAGCGCGCGCTGGACGGCAGGTGTTCCACGCGGCGACCGCGCTGA 1198
Qy 157 AAGAAGCAAGCTAGTCAAAAACGCGGTGCTGTACTTTGTGTGACTGCGGTAGGCGAAA 216
Db 1199 AGGATGCCAGAGGTACCTCCGGCGCGGTGCTCTGTGTGCCACCGCCATCGCGGAGA 1258
Qy 217 GTGTATTGAGACACAAACAAAGCGTTAAATTTGGCTGAGCAAAATCAATGCTCTGGGC 276
Db 1259 GCGTGTCCCGCGCCAGCAACAGGCGCTATCGGCTGCGCCGAGAGATCCCGTGAACGGCT 1318
Qy 277 GTTTTATCGTCGAGACATTTGGTTACAGGCTGTGGAACGAGAAC 321
Db 1319 GCTTCTACCGCAGGACATCGGCTACCGTGCCATCGCCCGCGGAGC 1363

RESULT 6
US-09-252-991A-10306
; Sequence 10306, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; CURRENT APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10306
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10306

Query Match 5.3%; Score 88.2; DB 4; Length 1383;
Best Local Similarity 56.8%; Pred. No. 3.2e-10;
Matches 162; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 37 TAGGTATCGTATTGGCTGACAGAGGATATCCAAAGGATTATCCAAAGGGCGATGAATCA 96
DB 957 TGGCGGTGGTACTGCGCGCGGCTATCCCGGCGACTACGCCAAGGGCGAGGTCTATCG 1016

QY 97 GCGGATTGCTTAAAGTGGCGGTCAAAAACGAGAAAGTTTCTTAGCGGGTGTGCGAGAAC 156
DB 1017 AAGGCTGCGCAGAGGCGCGCTGCGACGCAAGGTGTTCACGCGCGCACCGCGCTGA 1076

QY 157 AGAAGCGCAAGCTAGTACAAACGCGGCTGCTACTTTGTGTGACTGCTTAGGCGGAA 216
DB 1077 AGATGGCAGAGGTCACCTCCGCGCGCGGTCTGTGTGCGCCCGCTGCTGCGCGGCGAGA 1136

QY 217 GTGTATTTGAAGCACAAACAAAGCGTCAAAATTTGGCTGAGCAAAATTCATGCTGGGC 276
DB 1137 GCGTGTCCGCGCCAGCAACAGGCTATCGCCTGCGCGGAGAGATCCGCTGGAACGGCT 1196

QY 277 GTTTTATCGTCGACATGTTACAGGCTGTGGAACGAGAAC 321
DB 1197 GCTTCTACCGCAGGACATCGGCTACCGTCATCGCCCGCGGAGC 1241

RESULT 7
US-09-252-991A-10749/c
; Sequence 10749, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10749
; LENGTH: 1410
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10749

Query Match 5.3%; Score 88.2; DB 4; Length 1410;
Best Local Similarity 56.8%; Pred. No. 3.2e-10;
Matches 162; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 37 TAGGTATCGTATTGGCTGACAGAGGATATCCAAAGGATTATCCAAAGGGCGATGAATCA 96
DB 365 TGGCGGTGGTACTGCGCGCGGCTATCCCGGCGACTACGCCAAGGGCGAGGTCTATCG 306

QY 97 GCGGATTGCTTAAAGTGGCGGTCAAAAACGAGAAAGTTTCTTAGCGGGTGTGCGAGAAC 156
DB 305 AAGGCTGCGCAGAGGCGCGCTGCGACGCAAGGTGTTCACGCGCGCACCGCGCTGA 246

QY 157 AGAAGCGCAAGCTAGTACAAACGCGGCTGCTACTTTGTGTGACTGCTTAGGCGGAA 216
DB 245 AGGATGGCCAGAGGTCACTCCGCGCGCGGTGCTCTGTGCGCACCGCATCGCGGAGA 186

QY 217 GTGTATTTGAAGCACAAACAAAGCGTCAAAATTTGGCTGAGCAAAATTCATGCTGGGC 276
DB 185 GCGTGTCCCGCGCCAGCAACAGGCTATCGCCTGCGCGGAGAGATCCCGCTGGAACGGCT 126

QY 277 GTTTTATCGTCGACATGTTACAGGCTGTGGAACGAGAAC 321
DB 125 GCTTCTACCGCAGGACATCGGCTACCGTCATCGCCCGCGGAGC 81

RESULT 8

US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

Query Match 4.5%; Score 74; DB 4; Length 1141;
Best Local Similarity 11.0%; Pred. No. 3.7e-07;
Matches 110; Conservative 400; Mismatches 474; Indels 13; Gaps 4;

QY 358 GATAAAAAATATTGTACAGGGTAGAATTGTAATTTTCTTAGGATTTAGGATTTGTTAGGG 417
DB 56 SRKWTWARMYCKYRRWYNNKSRWKWYKKWYBCANNTSBRYHARRWKDMKTAYBMTWT 115

QY 418 CAACGTTTACGATTCGTCGACAAATAAATTAGAAATATTATTTTGTACTTATGAGGT 477
DB 116 NKWKGTRHRYRWRAABDTVDHHYVTAMNNAAWTTCMDKDDKTRTWWKNNNTGW 175

QY 478 TATATCAACTTATGCGACAATTTGTCATCGTAGTATATATCTGCAATTTCTTTTCAT 537
DB 176 DDDTKYHMNNNGCBTVTWYKTDWSBRYKMYGMBWKNWSYDVYIYVWVWDDMCK 235

QY 538 TAGAAGTTATTATAGAAAAATTATTTAATTTTCTAGTATTGAGAGATACACTGAAGTT 537
DB 236 RKVRVRVTRGRMRYVAVBTAHRRYNNNGWTBMAVYRRWTMNNNNNAKMKCKAYW 295

QY 598 ATTGTCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 657
DB 296 GWNRAVNSTCTTWSKTTTQVTSWANNCRAGDANKDKHKKWKSAAAGVYNNNNNNW 355

QY 658 TATTAGTTGGCGCTTATTGCAATAAGTATTGTTGTTAATAATGATGATGATGATGATGAT 716
DB 356 TYKKAHBAWDMVWHSAAWKWHAHAHYSRKKWTBYKRTKTVNNNNNNNNNNNNNNNNNN 415

Qy	717	TACAACTCTTGGATTGGACCTGTTAAATTACTCAGTTCGATTTAAAGAAATTAATGAGATA	776
Db	416	KMDMDBGTYNNNNNGGRTYYGTTKKNOMWYYKWKANNCKRWDHDKTCTHNNTTWKKM	475
Qy	777	ACAAATGCTGGCTTAAACAATGATAGATAAAATTCATATCATCATGTTTATTGTGTTTATT	836
Db	476	KTYWNNCYWKSMTNGKSHRBAAAVYTWYMWWRYYA---HANNNWWDYWKACTWYKYB	531
Qy	837	GAAGTGTGCTGTGTTTTAAGTTTCATAAAAAGAAAGTATATAAACTTCTCTGG	896
Db	532	VCSKWNVNAAYTKSSWNYTSRYRWKTNNSWRWSDTRSMGRANNYAPABHYGVKWN	591
Qy	897	ATTTTTGACCTTATTTTTTATGCTGATGATGTTGTTTGTTCGAGCGTATACAACA	956
Db	592	RWBWB----SHTWBHBAGAAHYWMMWYBAKCHCMKAWYKAKYAGAGSGNNNNNNNN	647
Qy	957	AAATCCCATGAGCGTTTTATTTCACCTTAACACACTGTTTATTCTCGATTAAATCCAA	1016
Db	648	NNNNNNATCARDDYYAASRWYAMANAkWYYKBAANNAYYTHANNWGWGNWATDTRT	707
Qy	1017	TTATCGTTGGTTATTTTATAGGACGAATTTGTTCCCTTATGAGATATTTTTCTTATCT	1076
Db	708	MWKNNNNNAGTWKNNNNNAKNAASAKNYAAAAVKAAKCHKWRWANKWVRGHADAAB	767
Qy	1077	ATTCTCTTTATCATAAATCTAAGCCCTATGAAATCGGGCTCTCCGVAATAATWARGAATA	1136
Db	768	TTDKBNAGATYKYTTTTNNNTYRGVTVNTAARDGWANNNNNNNNNNNNNGWSMDWTVW	827
Qy	1137	TAATTTTAAATTAATGGGGGAAGTGCACCTCAAGTCATTTTAGTGCTTT----TGTTA	1192
Db	828	AYANYGTNNNNNNNNAAYAWTKWYTTTDDRWBAYTNNNNNNRMYVYGAADDYAYY	887
Qy	1193	CGGAGAAAAACATCTCTTTTTTAGATAGCTTAAATATATAATCAGGAGCTCTTGTTG	1252
Db	888	MSDTCDAWKKWDATKMNNAATTYNGRTAWRTNNNNNNNTMTKYBYHAAWNNNNNNNG	947
Qy	1253	TAAAACTTATTCAGGAGGAAAGCTAAACAGCAATTTCTTTACCAATGTTTTTAAATGCA	1312
Db	948	AHTWVVCATKTKGCMWNCCTTCRKYKNNCTWYTWMTTTRTTWYAAATRWKTNATGSM	1007
Qy	1313	YCCTTAYCAAATGGAATACAACAGATAGCTAAAGGA	1349
Db	1008	RCNATGKNNNYTGWKTRWTAHYMATRWKMAWKVM	1044

```

RESULT 9
US-09-543-681A-3189
; Sequence 3189, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3189
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-3189

```

Qy 61 GATATCCAAAGGATTTATCGCAAGGCGATGAATCAGCGGATTGCCTTAAAGTGGCGTCA 120
 Db 1028 GCTATCCTGATACCTACCGCCAAATGATATTTAGGGACTTAACAGCAACATCTTCTT 1087
 Qy 121 AAAACGAGAAAGTTTCTTAGCGGGTGTGCGAGAACAG---AAGGCAAGCTAGTCACAA 177
 Db 1088 CAACAGCTAAGGTATTTCAAGCAGGAACAATAATTACGGCTCAAGGTGAAGTTGTTACAG 1147
 Qy 178 ACGGCGTCTGTGTACTTTTGTGTGACTGCGTTAGCGGAAAGTGATTTTGAAGCAACAACAA 237
 Db 1148 CAGGAGCCGAGTATTATGTGCAACTGCTTTAGGGGAAGATATCGAACAGCTCAAAAGA 1207
 Qy 238 AAGGTTAAATTTGGCTGACCAATTCATGCTGTGGCGGTTTTTATCTGTCGAGACATTTG 297
 Db 1208 ATGCGTATGCTTTTAGCCAAAAGTATTCACGTGATGGCTGTTTCTATCGCCATGATATCG 1267
 Qy 298 GTTACAGGCTGTGGAACGAGAACAA 323
 Db 1268 GCTATCGTGCCATTGCGCGTTTAAAN 1293
 RESULT 10
 US-10-204-708-2
 ; Sequence 2, Application US/10204708
 ; Patent NO. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/039971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 2
 ; LENGTH: 10467
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-2

[illegible]

RESULT 11
US-10-204-708-39
; Sequence 39, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIENPBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 39
; LENGTH: 19513
; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-39

	Query Match	4.3%	Score 71.6;	DB 4;	Length 19513;
	Best Local Similarity	47.4%	Pred. No. 2.6e-06;		
	Matches 287;	Conservative	0;	Mismatches 309;	Indels 10; Gaps 2;
Qy	337	AAATCTTTGTTGAATTTAATTAGATAAAAAATATTTGTACAGGTAGAAATTTGTAATTTTCCTA	396		
Db	8938	AAATATTATTGAATTTATATATATTTTAAATGGGTGAATTTGTATAGTATGTGAATTTATAGT	8997		
Qy	397	GGATTTAGGAATTTTGTAGGCAACGTTTACGATTGCTCTGACAAATTAATAGAAATTTAT	456		
Db	8998	TTATAAAGTAGTTTAGTAAATGAAGTTTTTATTTTCTTTTATAGGAAATTTAGTTTTTTA	9057		
Qy	457	ATTTTCTGTACTTTATCGAGGTTATATCAACTTATGCGACAATTTGTCTCATCGTAGTATTAT	516		
Db	9058	GTTTTTTTTTATAGAAATGATTAATTTAAAGTAAATAGATTTTATATATATATGTTTGTGATTAT	9117		
Qy	517	ATTCCTGCAA----TTCCTCTTTTCATTTAGAAGTTATTTTATAGAAAATTTATTTAAATTTCT	572		
Db	9118	AATAAAAAAGACGTTTATATATTTTGTATTATAAATAAAAAAGACGTTTATATATTTT	9177		
Qy	573	AGTATTGAGAGATACACTGAAAGTTATTTGTCGTGTTTGTGTTTGTGTTTGTGTTTGTGTT	632		
Db	9178	TGATTTAAAAAGTAAAAAGATTTTTTATATATTTTTTCTGTTTATTTTGTAAATTTTCTTATTA	9237		
Qy	633	TTTTCAAAATATAGAAATTTACAAGATATTAGTTGGCGCTTTATTTGCAATAAGTATTGTT	692		
Db	9238	TTTTTATTTTATATATTTTTTTTTTTTAAATATTTTGTGAAAAGTTTTTAAATTTGTAGAGAAGGT	9297		
Qy	693	GTTAATAATGTACATTTATGCAATATACCAATCTTGGATTTGGACCTGTTAATTTACTCACTT	752		
Db	9298	ATAAATTTGAGTATAGTGAATAGTTATGCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT	9357		
Qy	753	GCATTTAAAGAAATTAATAGATATACAAATGCTGGCTTAAACAATGATAGATAAATTCATA	812		
Db	9358	ATATTTTATTTGTTTATGTTTATTTTAAAGTATATGATATATATTTGT-----ATTTTTT	9411		
Qy	813	TATCCATTGTTATTTGCTTTTATTTGAAGTTGCTGCTGTTTTTAAAGTTTAAAGTTTTCATAAA	872		
Db	9412	GTTTTTTTTTTGTTGTTGAATTTATTTGAAAAGTTGTGTTTATATGATATTTTATTTTTTGAAA	9471		
Qy	873	AGAAAAGTATATAAACTTTCTTGGAATTTTTTGACTTTATTTTTTATGCTGTGATGATGAT	932		
Db	9472	ATTTTAGTAGTAAATTGTGTTT	9531		
Qy	933	GTTTTTT 938			
Db	9532	GTTATT 9537			

RESULT 12
 US-10-204-708-74
 ; Sequence 74, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7

[illegible]

Db 6160 ATTTTGGTAAAAAGTTTTTTTTCATGTTTCGGATTTTAGTTTTTTTATATTATTATTTTAA 6219
Qy 1105 TGAATCGGCTCTCCGVAABATTWARGAATATATATTTTAAATTAAT 1150
Db 6220 GATATTCGGGGGTTGGAGTTTATTTAGGGTTTATTTAGTTTAT 6265

Search completed: December 23, 2004, 19:48:59
Job time : 156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2004, 18:16:19 ; Search time 903 Seconds

(without alignments)
10211.434 Million cell updates/sec

Title: US-10-698-235-1

Perfect score: 1656

Sequence: 1 aagtaatatcacatagagg.....gggtgttkggaaraataacg 1656

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4105333 seqs, 2784095677 residues

Total number of hits satisfying chosen parameters: 8210666

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCT08_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	292.4	17.7	1830121	14	US-10-329-960-1
2	292.4	17.7	1830121	16	US-10-329-670-1
3	292.4	17.7	1830121	18	US-10-158-865-1
4	94	5.7	3673778	15	US-10-312-841-2
5	89.4	5.4	8056	18	US-10-473-126-386
6	87.2	5.3	19380	16	US-10-221-613-390
7	86.6	5.2	8079	16	US-10-240-589C-122
8	86.4	5.2	7814	16	US-10-221-714A-252
9	85.8	5.2	3673778	15	US-10-312-841-1
10	85.6	5.2	3683	18	US-10-473-126-339
11	85	5.1	1105	18	US-10-425-115-173590
12	84.6	5.1	1243	18	US-10-425-115-172717

13	84.2	5.1	6255	15	US-10-311-455-933	Sequence 933, App
14	83.8	5.1	8056	18	US-10-473-126-240	Sequence 240, App
15	83.6	5.0	8776	16	US-10-257-166-150	Sequence 150, App
16	83.4	5.0	5822	15	US-10-311-455-1069	Sequence 1069, App
17	83.2	5.0	1062	18	US-10-425-115-120013	Sequence 120013, App
18	83.2	5.0	7892	16	US-10-257-166-138	Sequence 138, App
19	82.8	5.0	6419	15	US-10-311-455-240	Sequence 240, App
20	82.8	5.0	113515	15	US-10-311-455-2147	Sequence 2147, App
21	82.2	5.0	8056	18	US-10-473-126-386	Sequence 386, App
22	81.8	4.9	6668	15	US-10-311-455-1670	Sequence 1670, App
23	81.4	4.9	17848	14	US-10-239-676-28	Sequence 28, App1
24	81.4	4.9	17848	15	US-10-240-453-38	Sequence 38, App1
25	81.4	4.9	17848	16	US-10-257-166-58	Sequence 58, App1
26	81.2	4.9	3683	18	US-10-473-126-193	Sequence 193, App
27	81.2	4.9	14147	15	US-10-172-086-51	Sequence 51, App1
28	81.2	4.9	14147	16	US-10-221-714A-469	Sequence 469, App
29	81.2	4.9	14147	17	US-10-311-507-39	Sequence 39, App1
30	81.2	4.9	14147	18	US-10-480-846-51	Sequence 51, App1
31	81	4.9	1081	18	US-10-425-115-16756	Sequence 16756, A
32	81	4.9	158001	16	US-10-211-179-11	GENERAL INFORMATION
33	80.6	4.9	921	18	US-10-425-115-38710	Sequence 38710, A
34	80.6	4.9	15548	15	US-10-311-455-2128	Sequence 2128, App
35	80.4	4.9	6292	16	US-10-221-714A-461	Sequence 461, App
36	80	4.8	7442	16	US-10-221-714A-409	Sequence 409, App
37	79.6	4.8	12405	14	US-10-239-676-35	Sequence 35, App1
38	79.6	4.8	12405	15	US-10-240-453-43	Sequence 43, App1
39	79.6	4.8	12405	16	US-10-221-613-101	Sequence 101, App
40	79.2	4.8	5979	14	US-10-239-676-17	Sequence 17, App1
41	79.2	4.8	5979	15	US-10-240-453-25	Sequence 25, App1
42	79.2	4.8	83391	17	US-10-433-793-123	Sequence 123, App
43	79	4.8	5452	15	US-10-311-455-1122	Sequence 1122, App
44	79	4.8	7498	15	US-10-311-455-230	Sequence 230, App
45	79	4.8	18683	15	US-10-311-455-286	Sequence 286, App

ALIGNMENTS

RESULT 1
US-10-329-960-1
; Sequence 1, Application US/10329960
; Publication No. US20030099277A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragment of Invention: Thereof, and Uses Thereof
; FILE REFERENCE: P186F1
; CURRENT APPLICATION NUMBER: US/10329,960
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c


```
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

;
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)

Query Match 17.7%; Score 292.4; DB 16; Length 1830121;
Best Local Similarity 85.1%; Pred. No. 5.6e-43;
Matches 336; Conservative 1; Mismatches 57; Indels 1; Gaps 1;

QY 8 ATCAATAGGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCAGAGGGATATCC 67
Db 941919 ATATCACAATGGATCTCGGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGGATATCC 941977
QY 68 AAAGGATTATCGCAAAAGCGGATGAAATCAGCGGATTGCTTAAAGTGCCTGCGGTCAAAAACGA 127
Db 941978 AAAGATTATCGCAAAAGCGGATGAAATCAGCGGATTGCTTAAAGTGCCTGCGGTCAAAAACGA 942037
QY 128 GAAAGTTTTCTTTCGCGGTGTCGAGACACAGAGGCAAGCTAGTCACAAAACGGCGTGC 187
Db 942038 GAAAGTTTTCTTTCGCGGTGTCGAGACACAGAGGCAAGCTAGTCACAAAACGGCGTGC 942097
QY 188 TGTACTTTGTGTGACTGCGTTAGCGGAAAGTGTTTGAAGCACAACAAAAGCGTTAAA 247
Db 942098 TGTACTTTGTGTGACTGCGTTAGCGGAAAGTGTTTGAAGCACAACAAAAGCGTTAAA 942157
QY 248 ATTGGCTGAGCAAAATTCATAGTCTGCGCGTGTTCGTCGAGACATTTGGTTACAGGCG 307
Db 942158 ATTGGCTGAGCAAAATTCATAGTCTGCGCGTGTTCGTCGAGACATTTGGTTACAGGCG 942217
QY 308 TGTGGAACGAGACAAAGCAAAATAGTTAGAAATCTTGTGTAATTTAATTAGATATAAAAT 367
Db 942218 TGTGGAACGAGATTCGCAAAAATAATCCATATCATATTATGATGAATTAATTTTAAAT 942277
QY 368 ATTGTACAGGAGTAAATTTGTTATTTTCCTAGGATTT 402
Db 942278 GATGAAATTTAGGATGAAATTTTAATTAATTT 942312

RESULT 3
```



```
us-10-158-865-1
; Sequence 1, Application US/10158865
; Publication NO. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; FILE REFERENCE: PB186P2C1D1
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 10/158,865
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-04-21
; PRIOR APPLICATION NUMBER: US 08/426,787
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)
```

```
Query Match 17.7%; Score 292.4; DB 18; Length 1830121;
Best Local Similarity 85.1%; Pred. No. 5.6e-43;
Matches 336; Conservative 1; Mismatches 57; Indels 1; Gaps 1;

QY 8 ATCAATAGSGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCAGAGGGATACC 67
DB 941919 ATATCAATAGGGATTCGCGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGGATACC 941977
QY 68 AAAGGATTATCGCAAGGGCGATGAATCAGCGGATTGCCTAAAGGCGGTCAAAAACGA 127
DB 941978 AAAAGATTATCGCAAGGGCGATGAATCAGCGGATTGCCTAAAGGCGGTCAAAAACGA 942037
QY 128 GAAAGTTTCTTAGCGGGTGTGCGAGAACGAAGGCAAGCTAGTACAAAACGGCGGTGCG 187
DB 942038 GAAAGTTTCTTAGCGGGTGTGCGAGAACGAAGGCAAGCTAGTACAAAACGGCGGTGCG 942097
QY 188 TGACTCTTGTGACTGCGTTAGCGGAAAGTGTTATTGAAGCACAACAAAAGCGGTAAA 247
```

```
Db 942098 TGFACCTTTGTGTGACTGCGTTAGGCGAAAGTGTATTGAAGCACAACAAAAGCGTTAAA 942157
QY 248 ATTGGCTGAGCAAAATTCAAATGCTCTGGGCGTTTTTATCTCGAGACATTGGTTACAGGCG 307
DB 942158 ATTGGCTGAGCAAAATTCATGCTCTGGGCGTTTTTATCTCGAGACATTGGTTACAGGCG 942217
QY 308 TGTGGAACGAGAACCAAGCAAAATAGTTAGAAATCTTGTTGAATTTAAATTAGATATAAAAT 367
DB 942218 TGTGGAACGAGAGTTGCAAAAATAATCCATATCATATGATGAATTAATTCATTTAAGT 942277
QY 368 ATTGTACAGGCTAGAAATTCGTTATTTCTTAGGATTT 402
DB 942278 GATGAAAATTTAGGAATGAAATTTTAATTAATTT 942312

RESULT 4
US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2 3673778
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2
```

```
Query Match 5.7%; Score 94; DB 15; Length 3673778;
Best Local Similarity 46.3%; Pred. No. 3.1e-06;
Matches 334; Conservative 3; Mismatches 384; Indels 1; Gaps 1;

QY 422 GTTACGATGCTCTGACAATAAATTAGAAATTAATTTTGTACTTTTACAGGTTATA 481
DB 257496 GTTTTATTAAGTTTTTTTATTTATTTTAAATTTTTTTTTTTTATTTATTTA 257555
QY 482 TCAACTTATGCGACAAATTTGTCATCGTAGTATTATTTCTGCAATTCCTTCATTAGA 541
DB 257556 TTAATTTATTTATTTTATTTAGTTAAATTTATGTATATAATTTATTTATTTAT 257615
QY 542 AGTTATTTATGAAAATTTATTAATTTCTAGTATTGAGAGATACACTGAAAGTTATT 601
DB 257616 TTATTTATTTATGATTATATACGTTTATTTAGTTATGATGAAGGTTTATTTATT 257675
QY 602 GTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 661
DB 257676 TGTAAATTTTTTTTGTGTTTAAATTTATTTTAAATTTATTTTAAATTTATTTAT 257735
QY 662 AGTTGCGGCTTTATTTGCAATAAGTATTGTTGTTTAAATGTTACATATTGACAGTATAC 721
DB 257736 TAATTTATTTTAAATTTTATTTAGTTATTTATTTATTTATTTATTTATTTATTT 257795
QY 722 ATCTTGGATTGGACCTGTAAATTTACTCTGCAATTTAAAGAAAATAATGAGATAACAA 781
DB 257796 ATTTATGATATAGAGATGATGATTTATTTAGTTATTTATTTATTTATTTATTTAT 257855
QY 782 TGCTGGCTTAACAATGATAGATAAATTCATATATCCATTGTTATTTGTTTATTTGAAAGT 841
DB 257856 TAAATTTAAATTTTATTTTAAATATATTTATTTATTTATTTTATTTATTTATTTAT 257915
QY 842 TGCTGCTGTTTTTAAGTTTTAAGTTTCATATAAAGAAAAGTATATAAATCTTCTTGATTT 901
DB 257916 TTATTCGATTAGTTATTT-ATTTATATAAGTTTATTTATTTATTTATTTATTTGTT 257974
```

Qy	902	TGACCTTATTTTATGCTGTGATGATGATGCTTTTGTTCGAGCGGTATACACAAATC	961
Db	25795	TTTATTTATTTGTAAATTTTCTGTTTATTTAGTTATTTTATACGTATATT	258034
Qy	962	CCATGAGCGTTTATTTTCCACTTAACACGTGTTTATTTCTCGATTAAAAATCAAATATTATC	1021
Db	258035	TATTTATTTATTTATTTATTAATGTTTATTTTATTTAAATGATATGAAGTTTTTAGTTATT	258094
Qy	1022	GTTGGGTTATTTTATAGGACGAATGTTCTCTATGAGATATTTCTTTATCTAAATATCC	1081
Db	258095	ATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA	258154
Qy	1082	TCTTTATCATAAATCTAAGCCCTATGAAATCGGGCTCTCGVAAATTTMARGAATATAATT	1141
Db	258155	AGTTAGTTAAATTTTTTTTGAATATTTGTTGGATAAGGAGAAATAAGAGAGATATAAAG	258214
Qy	1142	TT 1143	
Db	258215	TT 258216	

RESULT 5

US-10-473-126-386

; Sequence 386, Application US/10473126

; Publication No. US20040234973A1

; GENERAL INFORMATION:

; APPLICANT: Epigenomics AG

; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic proliferative disorders

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/473,126

; CURRENT FILING DATE: 2003-09-26

; NUMBER OF SEQ ID NOS: 1258

; SEQ ID NO 386

; LENGTH: 8056

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-473-126-386

Query Match	5.4%;	Score 89.4;	DB 18;	Length 8056;
Best Local Similarity	44.1%;	Pred. No. 2.1e-06;		
Matches	587;	Conservative 3;	Mismatches 729;	Indels 12; Gaps 5;
Qy	238	AAGCGTTAAAATTCGGCTCAGCAAAATCAATCGTCTGGCGGTTTTATATCGTCGACACATTG	297	
Db	752	AAATTTAAAATTTAATTAATAATTAAAAATTATATTTTAATTTATATATATTTAAAAATAAAATTT	811	
Qy	298	GTTACAGGCGTGTGGAAACGGAACAAGCAAAATAGTTAGAAATCTTGTTGAATTTAAATTA	357	
Db	812	ATAAATGATTATTTAAAAATAATTTTATTAATAAAAAAATGTTTTTTTATTTTATTTTATG	871	
Qy	358	GATAAAAAATATTCGTACAGGTAGAAATGTATTTTCTTAGAATTTAGGATTTGCTTAGGG	417	
Db	872	TATAAAAAATTTGTAAATTATTTAAATTTATATTTAAATATTTTGTGTTTGTGTAATA	931	
Qy	418	CAACGTTTTACGATTGCTCTGACAATAAAATAGAAATTAATTTTGTGTTACTTTATGAGGT	477	
Db	932	AAATTTTTTTAAAAATAAAATTTTTTTTATTTTAAATGAATAAAATATATATATAA	991	
Qy	478	TATATCAACTTATCGGCAAAATTCGTCAATCGTAGTATATATCTCGCAATCTCTCTTCAT	537	
Db	992	TTTTTTTTAAAAATAATTTATTTTAAAAAATTTTTTTTATTTATTTTAAATTTTGAATTTAAAT	1051	
Qy	538	TAGAAGTTATTTTATAGAAAAATTAATTTAAATAATCTCTAGTATTGACAGATACACTCAAAAGTT	597	
Db	1052	TATGTAATAATAAAAAAATTTTGTGTTTAAATTTGTTTTTTTTTAAATGTTAT	1111	
Qy	598	ATTGTCGTGTTGTTGTTGTTGTTGTTGTTGTTGTTTTCACAAATATAGA---ATTACAA	654	
Db	1112	ATAATTCGTTTTTAATAATTTTTTTTGTGATTTGTGAAATTTTAAAAATAATAATTTTTTGAT	1171	

RESULT 6
US-10-221-613-390
; Sequence 390, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian

Qy	655	GAATATTAGTTGGCGCTTTATTTTGCAATAAGTATTGTGTTAAATAATGTACATTATGCAG	714
Db	1172	TATTATTAAATTTATTTTAAAAATTTTAAATTTATTTAAATTTAAATTTTAAAAATAAAAT	1231
Qy	715	TATACCAAATCTTGGATTGGACCTGTTAAATTTACTCAGCTTGCATTTAAAGAANAATTAATGAGA	774
Db	1232	AATTTTTTTGAAATAAAAAAAATGTTATGTAATTTATTTAAATATATAAAAAATGTAATATT	1291
Qy	775	TAACAAAATGCTGGCTTAAACAATGATAGATAAAATTCATATCAATTCGATTGTTATTTGGTTTAT	834
Db	1292	TTTTTAAAAATTAATAAATTAATAAAAAATTTATTAATTTTAAATAAAAAATAATTTAA	1351
Qy	835	TTGAAGTGTGCTGTTTTTTAAAGTTTAAGTTTCATAAAAAAGAAAAGTATATAAACTTTCTT	894
Db	1352	ATTTTAAATTTTATTAAAAAATTTATTTTTTTTTTATAAAAATAAATTCATTTTTTTTTTT	1411
Qy	895	GGATTTTTTGACTTTATTTTTTATGCTGTGATGATGATGTTTGTGTTTCGAGCGTATACAA	954
Db	1412	TATTTTTTTTATTTTTTTTTTTTTAAAAAAAATAAAAAATTTATTTTTTTTAAATAAAATTTAT	1471
Qy	955	CAAAATCCCATGAGCGTTTTTATTTTCACTAACACTGTTTATCTCGATTAAAAATCCAAAT	1014
Db	1472	ATTAAAAATAATTAATTATTAATTAATTAATA-TAAATATAAATAATAAATATGTTAAA	1530
Qy	1015	ATTTATCGTTGGGTATTTTATAGGACGAATGTCTCTTAAGATAATTTCTTTTATCTA	1074
Db	1531	AAAAATTTTAAATAAAAAATTTATTTTTTAAAAAATTAATTTAAATTTAAATTTTAAATTT	1590
Qy	1075	ATATTCCTCTTTATCATAAATCTTAAGCCATGAATCGGCCTCTCCGVAAAAATMARGAA	1134
Db	1591	ATATTTATAATTTTAATATTTTAAAAAATTTGAATAAATGAATTT--GTAAAAATTAATAAAA	1648
Qy	1135	TATAATTTTAAATTAATGGGGGAAAGTCGACCTCAAGTCATTTTAGTGCCTTTTGGTTACG	1194
Db	1649	AATTAATTTTAAATTAATGATAAATTTTATTTTATTAATAATTAATAAATTAATTTAA	1708
Qy	1195	GGAGAAAA---ACATCTCTTTTTTATAGATAGCTTAAAAATAATAATCAGGAGCTCTTGTT-	1250
Db	1709	TATAATAATAATATATTTATAATTTTATTTTAAAAAATTAATAAATAAAAAATTTTTTA	1768
Qy	1251	--GGTAAAACTTATCAGGAGGAAAGCTAACAGCAATTTCTTTTACCNAATGTTTTTAAATG	1308
Db	1769	AATTAATAAATAATAAATTAATTTATGTTATTAATTTTAAATTAATAAATAAATAAATTT	1828
Qy	1309	CAATYCCCTTAYCCAAATGGAATACAAGATAGCTAAAGAGAGATACGAATTTTATTTAAATTT	1368
Db	1829	TTGTTTAAAAAATAAAAAAATAAAAAAATAAATTAATTTAAATTTATTTTATTTT	1888
Qy	1369	TAGCGAAGAGCAAGGCTTTCAGACATATTTTTTATTCAGCTCAAGCTAGGGATGATATGC	1428
Db	1889	TTTTATTATAAAAAATAAAAAATTTTATAAAAAAATAAATAAATTAATAATAAATAAATAA	1948
Qy	1429	ATATGATCAATTTTTTAGGAGGAGCTTGGATTCGATTCGATTTCCGTTTTCCAGATAATCAAG	1488
Db	1949	AAAAATATAATTTTTTAAAAATAAAAAATTAATAATTAATTTTATTTAAAAATTTAATAAAA	2008
Qy	1489	GGTATTCCTTTAAGAGATTCCAATGCCCTGAATAAATTAATTCCTCGCTTTTAAAAAATTTA	1548
Db	2009	TTTTTAAATAATTTAAAAATAAATAATTTTTTAAAAATTTTTTTATTTTAAAAAATTTTA	2068
Qy	1549	ATTTAGATAAT	1559
Db	2069	TTAAAAATATT	2079

```
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221.613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 390
; LENGTH: 19380
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-390

Query Match 5.3%; Score 87.2; DB 16; Length 19380;
Best Local Similarity 43.6%; Pred. No. 7.6e-06;
Matches 386; Conservative 1; Mismatches 499; Indels 0; Gaps 0;

Qy 348 AATTAAATAGATAAAAAATTTGACAGGAGTAAATTTCTCTAGGATTTAGGAT 407
Db |||||
Qy 11826 AATATTAAATGAATGATATATGATAGAAATGATTTGTTAAATAATAAAA 11885
Db |||||
Qy 408 TTCTGTAGGCAACGTTTACGATTCCTGCAATAAATAGAAATTAATTTTGTAC 467
Db |||||
Qy 11886 ATGAATAATTATATATGATATATTTAGTTTAAATAAAATATTTGTTATTTGAATA 11945
Db |||||
Qy 468 TTTATGAGTTATATCACTATATCGCAATTTCTCATCGTAGTATATATCTTGCATTT 527
Db |||||
Qy 11946 TTTTATGATATATAAATGAATTAAGATATATTTATTCGTAATATATAAAAAATTTT 12005
Db |||||
Qy 528 CTTCCTTCATTAGAGTTATTTATAGAAAAATTTTAAATATTTCTAGTATTTGAGAGATAC 587
Db |||||
Qy 12006 TTTTAAATATTTTATTTATGATGGAATTTTGTAGTTTAAATGTTTAAAGTATTTT 12065
Db |||||
Qy 588 ACTGAAAGTTATTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 647
Db |||||
Qy 12066 TTTATTTTATAAATTTATTTTATTTTAAATGGTTGTTGTTGTTGTTGTTGTTGTTGTT 12125
Db |||||
Qy 648 ATTACAAGATATTTAGTTGGCGCTTTTATTTGCAATAAGTATTTGTTTAAATAATGTACAT 707
Db |||||
Qy 12126 AGTTATGGAATATATAAGGATAGATTAATTTTAAATTTGAAATTTATTTTAAATTTTGT 12185
Db |||||
Qy 708 TATGAGTATACCAATCTTGGATTTGGACCTGTTAAATTAATCTCACTTGCATTTTAAAGAAAT 767
Db |||||
Qy 12186 TAGTTTTTTTATAGATATATAATTTGATGTTTAAATTAAGATGTTTTTTTTTATAGT 12245
Db |||||
Qy 768 AATGAGATAACAATGCTGGCTTAAACATGATAGATAAAATTCATATATCCATTTGTTATTT 827
Db |||||
Qy 12246 TATTATATATATATTTATATATATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 12305
Db |||||
Qy 828 GGTATTATTTGAGTTGCTGCTGTTTTTAAAGTTTAAAGTTTCAATAAAGAAAGATATAAAA 887
Db |||||
Qy 12306 AAGAAATAGAAATTTTTTTTATAGATGATATAGATAAATGATATTTGAAATTTAAATGAT 12365
Db |||||
Qy 888 CTTCCTTGGATTTTGAATTTTATTTTATGCTGTGATGATGATGTTTTTTTGTTCGAGCG 947
Db |||||
Qy 12366 GTATTTTTTGAATTTAAATTTTGTATATAGGAATTTATTTAGGTTAAATTTTATATTTATG 12425
Db |||||
Qy 948 TATACACAAATCCCATGAGGTTTTTATTTTCACTTAACACTGTTTATCTCGAATATAAA 1007
Db |||||
Qy 12426 TGGAAATGATATATGGGAAAGAAATGTTTATAGTTAAATTTTGAATTTTGGAGTATTTTAT 12485
Db |||||
```

RESULT 7

```
US-10-240-589C-122
; Sequence 122, Application US/10240589C
; Publication No. US20040076956A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: DNA repair
; FILE REFERENCE: 5013.1008
; CURRENT APPLICATION NUMBER: US/10/240,589C
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03972
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 148
; SEQ ID NO 122
; LENGTH: 8079
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-589C-122
```

```
Query Match 5.2%; Score 86.6; DB 16; Length 8079;
Best Local Similarity 46.4%; Pred. No. 7e-06;
Matches 352; Conservative 0; Mismatches 404; Indels 3; Gaps 2;

Qy 333 TTAGAAATCTGTGTGAATTTAATAGATAAAAAATATTGTACAGGGTAGAATTTGATTTT 392
Db |||||
Qy 4841 TTATATATTTTGTGTTTTTTTGTGTTATAGTTTGTGTTTTTATTTATTTAGTAGTTGAGTG 4900
Db |||||
Qy 393 CCTAGGATTTAGGATTTGTTAGGCAACGTTTACGATTCCTCTGACAAATAAATAGNAT 452
Db |||||
Qy 4901 ATTAGGTTATGATTTTTTTTTTGTTTAAATTTTTTTAGTGGTTTTTATTTTATTTGGGT 4960
Db |||||
Qy 453 TATTATTTTGTACTTTATGAGGTTATATCAACTTATGCGACAAT-TTGTCACTCGTAGT 511
Db |||||
Qy 4961 AAAAAATGAAAAATTTTATAGTAGTTTATAGGTTTATAGATGATTTTCTGTTATTTTATTT 5020
Db |||||
Qy 512 ATTATATCTGCAATCTCTTTTCAATTTAGAAATTTATTTATAGAAAAATTTTAAATATTTTC 571
Db |||||
Qy 5021 TTTTAAATTTTAGTTTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTGGT 5080
Db |||||
Qy 572 TAGTATTTAGAGATACACTGAAAGTATTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 631
Db |||||
Qy 5081 TTGTTTTTAAAGTATGTTAGTTATGTTTTTTGTTTTTATAGATTTTGTATTTTGTGTTTTTT 5140
Db |||||
```


QY	335	AGAACTCTGTGTAATTAATAGATAAAAAATATCTACAGGCTAGAAATGTAATTTCC	394
Db	1058460	AGTATTTAGTTAAAGTATTTGGTGATAGATTTGGGAATAGAAATTTGGATTTTAGATT	1058519
QY	395	TAGGAATTTAGGAATTTTCTTAGGGCAACGTTTACGATTTGCTCTGCACATAAAATTAGAATTA	454
Db	1058520	TTTAGTTTAGGGTTTTTTTTTATTGTAATTTGGGGCTTTTTCGAATTTTTTTTTTTTTTTT	1058579
QY	455	TTATTTTGTGTTACTTTATGAGGTTATATCAACTATACGACAAATTTGTCATCGTAGTAAT	514
Db	1058580	TTTATTTTGTGTAATTTTAAAGTTTCTTTGATTTTTTTTTTATGAAGGTTTTTTTTTAATTTTTT	1058639
QY	515	ATATTTCTGCAATCTCTCTTCATTTAGAAAGTTATTTATAGAAAAATTTTAAATATTTCTAG	574
Db	1058640	ATTTTTTTTTTATTTTATATATATAGGTTTAAAGTTTAAATATATTTT-TAGTTTGAG	1058698
QY	575	TATTTAGAGATACACTGAAAGTTATTTGCTGTTGTTGTTTGTGTTT-----GTT	624
Db	1058699	TATTTATTTTTTATTTTTTATTTTTTATTTTTTGTGTTTGTGTTTGTGTTTAGATGTTT	1058758
QY	625	TGTTTGTGTTTTTCAAATATAGAAATTAACAAGATATTTAGTTGGGCTTTATTTTGCATATA	684
Db	1058759	TTTTTATTTTGTCTGTTAAAGATGTTTTTATAGATGTACGTATGTTGTTTTTGTGTTTGA	1058818
QY	685	GTATTTGTGTTTAAATATGATACATTTATGCAGTATATACCAATCTTGGATTTGGACCTGTTAAAT	744
Db	1058819	ATTTTAGTGTTTATTTATAGGTTTTTGAAGATATATTTAAATTTTTATATCGGTATTT	1058878
QY	745	ACTCATTGTCATTTAAAGAAATTAATGAGATAACAATGCTGCTTTAAACAATGATAGATA	804
Db	1058879	AGTTAGTTTAGTGTGTTTATAATAAAATATTTTAGACGGGTGTTTTAAATTAATAGAAGTT	1058938
QY	805	AATTCATATATCCATTTGTTTATTTGGTTTATTTGAAGTTGCTGTGTTTTTAAAGTTTAAAGTT	864
Db	1058939	TATTTTTATAGTTTTCGAAGTTTGAGAAAGTTTAAAGTTTGGTGAAGGTTAGTTT	1058998
QY	865	TCATAAAAGAAAAGTATATAAACTTCTTGGATTTTGGATTTATTTTATGCTGTGA	924
Db	1058999	TATTTAGATTTTAGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTTATTTATTTT	1059058
QY	925	TGATGTATGTT 935	
Db	1059059	TTATTTTTTT 1059069	
RESULT 10			
US-10-473-126-339			
; Sequence 339, Application US/10473126			
; Publication No. US20040234973A1			
; GENERAL INFORMATION:			
; APPLICANT: Epigenomics AG			
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell			
; FILE REFERENCE:			
; CURRENT APPLICATION NUMBER: US/10/473,126			
; CURRENT FILING DATE: 2003-09-26			
; NUMBER OF SEQ ID NOS: 1258			
; SEQ ID NO 339			
; LENGTH: 3683			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)			
US-10-473-126-339			
Query Match 5.2%; Score 85.6; DB 18; Length 3683;			
Best Local Similarity 47.9%; Pred. No. 7.9e-06;			
Matches 312; Conservative 0; Mismatches 334; Indels 6; Gaps 2;			
QY	442	TAAATAGAAATTAATTTTGTGTTACTTTATGAGTTATATCACTTATATCGCAATTTG	501
Db	2724	TATTTTTTTTTTATTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTTGTGTTTTT	2783

QY	502	TCATCGTAGTATTATATCTCGCAATTTCTTTTCATTAGAAGTTATTTATAGAAAAATTAT	561
Db	2784	ATATATTTTATTTTATTTTAAATTTTGTGTTTTTTTTTATTTATTTATTTATTTGTTTT	2843
QY	562	TTAATAATTTCTAGTATTTAGAGATACACGAAAGTTATTTTGTCTGTTGTTGTTGTTTT	621
Db	2844	TTTTTTTTTATTTTGTGTTTTTGTGTTTTTTTTTTTTTTTTTTTTTATTTATTTT	2903
QY	622	GTTTGTGTTGTTTTTCAAATATAGAAATTTACAAGAATATAGTTGGCGCTTTATTTGCAA	681
Db	2904	TTTTTTTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTTT	2963
QY	682	TAAGTATTTGTTTAAATAATGTACATTATGCAGTATACCAATCTTGGATTTGGACCTGTTA	741
Db	2964	TTTTTATTT-TTTTTTATTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTT	3022
QY	742	ATTACTCACTTGCATTTAAAGAAATTAATGAGATAACAATGCTGCTTTAAACAATGATAG	801
Db	3023	TTTTTGTGTTTTTAAATTTTATGTTTTTATTTTGTGTTTTTTTTTTTTTTTTTATTT	3082
QY	802	ATAAATTCATATATCCATTTGTTATTTGGTTTTATTTGAAGTTGCTGTGTTTTTAAAGTTTAA	861
Db	3083	TTTATTTTATTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTGTGTTTTTT	3142
QY	862	GTTTTCATAAAAGAAAAGTATATAAACTTCTTGGATTTTTCGACTTTATTTTATGCTG	921
Db	3143	ATTTGTTTTTTTTTTTGTATGTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTATGTT	3202
QY	922	TGATGATGTATGTTTTTGTTCGAGCGTATACACAAATCCCAT-----GAGCGTTTTTAT	976
Db	3203	TATTTTTTTTTTATTTGTTTTTTTTTTTTTTTTTATGTTTATTTTTTTTTTATTTGTTTTT	3262
QY	977	TTCCACCTACACTGTTTATTTCTCGATTAATAAATCCCAATTTATCGTTGGGTTATTTTAT	1036
Db	3263	TTTTTTTATGTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTATTTTAAATTTTTTTT	3322
QY	1037	AGGACGAATGTTTCTTATGAGATATTTTCTTTATCTAATATTTCCCTCTTTAT	1088
Db	3323	TGTTTTTTTTTTTTTATTTTATTTTATTTTTTTTTTTTTTTTTTATTTTATTTGTTTT	3374

RESULT 11

US-10-425-115-173590
; Sequence 173590, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 173590
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1105)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_89895C.1
US-10-425-115-173590

Query Match 5.1%; Score 85; DB 18; Length 1105;
Best Local Similarity 46.4%; Pred. No. 6.5e-06;
Matches 277; Conservative 0; Mismatches 320; Indels 0; Gaps 0;

Db 1557 AAAAAAAAAATTTTATTTAAAAATTTTAAACGTATTT--ATATTTTATATTTATATATAA 1500
QY 1335 CAGATAGCTAAAGGAGATACGAATTTATTTAAATTAGCGAAAGAGCAAGCCTTTCAGACA 1394
Db 1499 TATTAATTTAATAATTAATTTATTTAATAATAATTTTAAAAAATAAATTTTATTTT 1440
QY 1395 TATTTTATTCAGCTCAAGCTAGGAGATGATATGCATAT--GATCAATTTTATAGGAGGA 1451
Db 1439 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1380
QY 1452 GCTCGATTCGATGATATTCGTTTTCAGATATATGAGGGTATTCCTTTAGAGATTCGAATG 1511
Db 1379 AATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1320
QY 1512 CCGTAAATAAATAATCTCTGCTCTTAAAAATAATTAATTTAGATAATGGTTATCATTTT 1571
Db 1319 TTTTATTAATTTATTAATTTTAAAAAATAATACGTTTATATATATTAATAATAATTAC 1260
QY 1572 GTTGTTTTACATCAGAGGGAGTCAATTCCTATGGGGCATATTAAGAATGAATAAAGN 1631
Db 1259 GTAAACGTTTTTTTATTCGAAAAATTTATTTATTTAAAAAATAAATAAATAAATAA 1200
QY 1632 AGNAGGTCGTTKGGAAAAATAA 1654
Db 1199 ATTAAAAATTTTAAAAATAAATAA 1177

RESULT 15
US-10-257-166-150
; Sequence 150, Application US/10257166
; Publication No. US20040023230A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
; TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
; FILE REFERENCE: 5013.1011
; CURRENT APPLICATION NUMBER: US/10/257,166
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/EP01/07470
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-06-29
; 2000-09-01
; 2000-09-01
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 150
; LENGTH: 8776
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-150

Query Match 5.0%; Score 83.6; DB 16; Length 8776;
Best Local Similarity 47.0%; Pred. No. 2.6e-05;
Matches 359; Conservative 0; Mismatches 399; Indels 6; Gaps 3;
QY 339 ATCTGTGGAAATTTAATAGATAAATAAATAATTCGTACAGGGTAGAATGTATTTTCCTAGG 398
Db 6685 ATATTTTATTTTGTAGATTTGTTTATTTGTTGGTATGTAGAAAATGTTATTTGGTTTTG 6744
QY 399 ATTTAGATTTTGTAGGCAAGCTTTACGATTCCTCACAATAAATTAGAAATTTATAT 458
Db 6745 TAGTTGATTTTGTATTTAGTAATAATATTTGAATTTGTTTATTTAGATTTTAAAGTTTTT 6804
QY 459 TTTTGTACTTATGAGGTTATATCAACTTATGCGCAATTTGTCATCGTAGTATTTATAT 518
Db 6805 AGTGGAGTTTTTAAAGTGTTTTATATATTTATTTGGAATTTATTTGGAATAATTTTGGCA 6864
QY 519 TCTGCAATTCCTCTTTTCATAGAGTTATTTATAGAAAATTTATTTAAATTTCTAGTATT 578

Db 6865 AATTTATTTTAAATATTTATTTTAAATATATTTTGGAAATTTATTTGAAATATTTTGTGTT 6924
QY 579 GAGAGATACACTGAAAGTTTATTTGTCGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTCA 638
Db 6925 TTTTGGTATA---AAATGTAATTTTAGGTTTATTTTGAATATTTTAAATTTTGGTTTTGGA 6981
QY 639 AAATATAGAAATTTACAAGAAATATTTAGTTGGCGCTTATTTTGGCAATAAGTATTTGTGTAAT 698
Db 6982 ATTAGTTATTTTAAAAAAGTTTGGTTTTTTTAGTCGAAAAATGCTATTAATAAGTTA 7041
QY 699 AATGTACATTTATGTCAGTATACCAATCTTGGATTGGACCTGTTAAATTTACCTCAGCTGCAT- 757
Db 7042 AGATTTGGTTTATTAAGTGTGTTTATTTGTTATTTGAATGTTATTTGTTATTTAGTTTATTA 7101
QY 758 -TAAAGAAATTAATGAGATACAAATGCTGGCTTAACAATGATAGATAAAATTCATATATC 816
Db 7102 GTGGAGAGAAGTAGGAGTATATGTTTTTTTCTGATATATATATAAATATATATATATAT 7161
QY 817 CATTTGTTATTTGGTTTATTTTGAAGTTGCTGCTGTTTTTAAAGTTTAAAGTTTCAATAAAGAA 876
Db 7162 TATGTTTATTTAAATTTTGTGTTTTTTTGTGTTATTTTATTTTATTTTAAATTTATTTAT 7221
QY 877 AAGTATATAAACTTTCTTGGATTTTGTGATTTTGTACTTTTATTTTATGCTGTGATGATG-TT 935
Db 7222 TTATTTTAGAAGGGTATGAATATATATGGATATATTTATTTTAAATTTTATAGGAT 7281
QY 936 TTTGTTTCGAGCGTATACAACAAAATCCCATGAGCGTTTTTATTTTCACCTAACACTGTTTAT 995
Db 7282 TTTGTTAGTTTTTTTTTTTTTTTGAATTTATGATTTTTTTTTTGTAAAAGTGATATATTTGG 7341
QY 996 TCTCGAATTAATAATCCAAATTTATTTATCGTTGGGTTATTTTATAGGACGAATTTGTTCTTAT 1055
Db 7342 TTTTATGATATATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 7401
QY 1056 GAGATATTTTCTTTATCTAATATCTCTCTTTATCATATCTAA 1099
Db 7402 ATTTTATTTTATTTATGATGTTTTTTTTTATTTATTTATGTTTTTAAA 7445

Search completed: December 23, 2004, 21:59:10
Job time : 914 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2004, 15:46:49 ; Search time 5397 Seconds
(without alignments)

11181.063 Million cell updates/sec

Title: US-10-698-235-1

Perfect score: 1656

Sequence: 1 aagtaataatcacatagagg.....gggtgttkggaaraataaacg 1656

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsa1:*

9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107.8	6.5	1254	9	AG349719 Mus muscu
2	103.8	6.3	2157	9	CL081966 CH216-165
3	103.2	6.2	1277	8	CC253231 CH261-180
4	102.8	6.2	1536	9	CL078538 CH216-151
5	102.6	6.2	1205	9	CL143963 ISB1-1230
6	102.4	6.2	1162	9	CL077122 CH216-143
7	102.2	6.2	1493	9	CL078589 CH216-151
8	102	6.2	1805	9	CL080711 CH216-159
9	101.4	6.1	1826	6	CF238805 AGENCOURT
10	101.2	6.1	1542	9	AG386981 Mus muscu
11	99.8	6.0	1599	9	CL083840 ISB1-2H14
12	99	6.0	1594	9	CL038406 CH216-46A
13	98.6	6.0	1981	9	CL082000 CH216-165
14	98.4	5.9	1344	9	CG748432 P042-2-C1
15	98.4	5.9	1632	9	CL082569 CH216-167
16	98	5.9	1297	9	CG758143 P053-3-B1
17	98	5.9	1784	9	CL081992 CH216-165
18	97.8	5.9	1101	8	AL069706 Drosophil
19	97.6	5.9	1206	8	BZ695529 SP_Ba006
20	97.6	5.9	1566	9	CG757757 P053-1-D0
21	97.2	5.9	1339	8	AQ897537 HS_3153_A
22	97	5.9	1638	9	AG347262 Mus muscu
23	96.8	5.8	1101	9	AL063921 Drosophil
24	96.6	5.8	1811	9	CG753732 P048-4-GO

C 25	96.4	5.8	1310	9	AG370846	Mus muscu
C 26	96.2	5.8	1512	9	CL082685	CH216-169
C 27	96	5.8	1392	9	CG757503	P052-4-CO
C 28	95.8	5.8	1885	2	BE420745	HM002.BO
C 29	95.4	5.8	1896	3	CR722884	Tetraodon
C 30	95	5.7	1324	9	AG376784	Mus muscu
C 31	94.8	5.7	994	9	CNS04NOJ	AL298972 Tetraodon
C 32	94.6	5.7	1251	9	AG332167	Mus muscu
C 33	94.6	5.7	1433	9	CG745119	P037-4-GO
C 34	94.2	5.7	1042	9	CL461494	SAIL.1148
C 35	93.8	5.7	1135	9	CNS033GQ	AL226115 Tetraodon
C 36	93.8	5.7	1407	9	AJ592026	Arabidops
C 37	93.4	5.6	938	9	CL509354	SAIL.811
C 38	93.4	5.6	1074	8	BZ696936	SP_Ba009
C 39	93.4	5.6	1289	9	AG347097	Mus muscu
C 40	93.4	5.6	1348	9	CG749499	P043-4-AO
C 41	93.4	5.6	1489	9	CL078612	CH216-151
C 42	93.2	5.6	1346	9	AG382256	Mus muscu
C 43	93.2	5.6	1355	9	AG346348	Mus muscu
C 44	93.2	5.6	1539	9	AG340947	Mus muscu
C 45	93.2	5.6	2142	3	CR730230	Tetraodon

ALIGNMENTS

RESULT 1	AG349719	Mus musculus molossinus	1254 bp	DNA	linear	GSS 02-JUN-2004
LOCUS	AG349719	Mus musculus molossinus	sequence.			
ACCESSION	AG349719					
VERSION	AG349719.1	GI:47923029				
KEYWORDS	GSS					
SOURCE	Mus musculus molossinus					
ORGANISM	Mus musculus molossinus					
REFERENCE	1					
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.					
TITLE	BAC end Sequences of Library MSMg01					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 1254)					
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.					
TITLE	Direct Submission					
JOURNAL	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan (E-mail:hattori@psc.riken.jp, URL:http://hgp.gsc.riken.go.jp/), 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel:81-45-503-9111, Fax:81-45-503-9170					
COMMENT	Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp					
PRIMERS	Sequencing : T7					
LIBRARY	Vector : pBACe3.6					
R.Site 1	: EcoRI					
R.Site 2	: EcoRI					
FEATURES	Location/Qualifiers					
source	1..1254					
	/organism="Mus musculus molossinus"					
	/mol_type="genomic DNA"					
	/sub_species="molossinus"					
	/db_xref="taxon:57486"					
	/clone="MSMg01-146B10.T7"					
	/sex="male"					
	/tissue_type="mixture of kidney and spleen"					
	/clone_lib="MSMg01 Mouse Male BAC Library"					

```
ORIGIN
Query Match      6.5%; Score 107.8; DB 9; Length 1254;
Best Local Similarity 47.4%; Pred. No. 5.4e-10;
Matches 390; Conservative 0; Mismatches 427; Indels 6; Gaps 2;

QY 279 TTTTATCGTCGAGACATTTGGTTACAGGCGCTGTGGAACGAGACAACGAAAATAGTTAGAA 338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 369 TTTTATTAAGATGATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTATATAA 428

QY 339 ATCTGTGTAATTAATTAAGATAAAAAATATGTACAGGGAGAAATTTGTATTTTCCTAGG 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 429 AAATTAATTTTATTCATTTTATATTAATTTTGTGTATATATATATATATTTTATATAT 488

QY 399 ATTTAGGATTTTGTAGGCAACCTTACGA-TTGGCTCTGACATAAATTAGAAATATTA 457
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 489 ATTTGTTATTTTATTTGTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 548

QY 458 TTTTGTGTACCTTATGAGGTTATATCAACTTATCGGACAATTTGTCATCGTAGTATTATA 517
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 549 TTTTATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATAT 608

QY 518 TTCTGCAATCTCTTCATTAAGAGTTATTTATAGAAAATATTTAATATTTCTAGTAT 577
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 609 TTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 668

QY 578 TGAGAGATACACTGAAAGTTATTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTTTC 637
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 669 TTATAAATTTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTTATTA 728

QY 638 AAAATATAGAAATTAACAAGATATTAGTTGGCGCTTTATTTGCAATAAGTATTGTGTAA 697
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 729 TTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 788

QY 698 TAATGTACATATCGAGTATACCAATCTTGGATTTGGACCTGTGTTAATTAATCACTGCAAT 757
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 789 TTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 848

QY 758 TAAAGAAATTAATGAGATAACAAATGCTGGCTTAACAATGATAGATAAATTCATATATCC 817
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 849 TAATATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTA 908

QY 818 ATT-----GTTATTTGGTTATTTGAAAGTTGCTGTGTTTAAAGTTTAAAGTTTCAATAAA 872
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 909 TTTAATTAATTTATTTTGTATTATATA-TTAAATTTATTTTATTTTATTTTATTTATTTT 968

QY 873 AGAAAGTATATAAATCTTCTGGATTTTGGACTTTATTTTATTTTATTTGCTGTGATGATAT 932
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 969 AATAAATTTATTTGATTTTATTTTATTTTATTTTATTTTATTTATTAATAATTTTATTTAT 1028

QY 933 GTTTTGTTCGAGCGTATACAACAAAATCCCATGAGCGTTTATTTCCACCTAACACTGTT 992
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1029 TATATTTTATTTAATATATATATTTTATTTATTTTATTTTATTTAATATATATATATTTT 1088

QY 993 TATTCGATTAATAATCCAAATTAATTAATCGTTGGTTATTTTATAGGACGAATTTGTCT 1052
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1089 TATTTATTTTATATATTTTATTTTATTTATTTATTTATTTTATTTTATTTTATAGTATA 1148

QY 1053 TATGAGATATTTCTTATCTAATATTCCTCTTTTATCATAAAT 1095
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1149 TATTAATTTTATTAATTTTATTAATTTATTTTATTTATTTAAT 1191

RESULT 2
LOCUS CL081966
DEFINITION CH216-165D13_Sp5.1 CH216 xenopus tropicalis genomic clone
ACCESSION CL081966
VERSION CL081966.1
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
```

```
REFERENCE
AUTHORS Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp5 atctgcggttcgatcct
Class: BAC ends
High quality sequence start: 341
High quality sequence stop: 412.
Location/Qualifiers
1. .2157
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-165D13"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match      6.3%; Score 103.8; DB 9; Length 2157;
Best Local Similarity 47.8%; Pred. No. 3e-09;
Matches 427; Conservative 0; Mismatches 458; Indels 9; Gaps 4;

QY 350 TTTAATTAGATAAAAAATATTTGACAGGAGTAATTTGTTATTTTCTCTAGGATTTAGGATTT 409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1148 TTTATTTATTTATAGTTTTTTTTTTTATTTATTTATTTATTTATTTTATTTTATTTT 1207

QY 410 TGTAGGGCAACGTTTACGATTTGCTCTGACATAAATTAGAAATTTATTTTGTACTT 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1208 TATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1267

QY 470 TATGAGGTATATCAACTTATCGGACA-ATTTGTCTAGTAGTATTATTTCTGCAATTC 528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1268 TTTTATGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 1327

QY 529 TTTCTTCATTAGAGTTATTTATAGAAAATTTTAAATTTCTAGTATTGAGAGATACA 588
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1328 TATATTTATTTTATTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTT 1387

QY 589 CTGAAAGTTATTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTTTCARAAATAGAA 648
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1388 TTATATTTTATATATATGTTTTTTTAAATTTGTTTATTTTATTTTATATATTA 1447

QY 649 TTACAAGATATTTAGTTGGCGCTTTATTTGCAATAAGTATTTGTTAATAATTTGACATT 708
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1448 TTTTATTTATTTTATATATGTTTATTTATTTATTTATTTATTTTATTTTATTTT 1507

QY 709 ATGCAGTATACCAATCTTGGATTTGGACCTGTTTAAATCTCACTTGCATTTAAAGAAATTA 768
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1508 TTAATTTTTTTTATTTATTTGTTTATTTTATTTATTTTATTTTATTTTATTTT 1567

QY 769 ATGAGATAACAATCTGGCTTAAACAATGATAGATAAATTCATATATCCATTGTTATTG 828
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1568 TTTTATTTTATTTATTTGTTTCATATATTTTATTTAAATTTTATTTATTTATTTA 1627

QY 829 GTTTATTTGAGTTCGCTGTGTTTATTTAAAGTTTTCATAAAAAGAAAGTATATAAAC 888
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1628 TTTTATTTTGTGTTTGTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1687

QY 889 TTTCTTG----GATTTTTCACCTTTATTTTATGCTGTGATGATGTTTGTTCGA 944
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

[illegible]

382	Qy	AATGTGATTTTCTCAGATTAGAGTTTTCGTAGGGCAACGTTTACGATTGCTCTGCAAA	441
1179	Db	TA'TTAAATATATATTTTTTATPAAATATATATAAAATTTTTTTTAAATTTTATTTATATATTA	1120
442	Qy	TAAATTAGAATATATTTTTTGTTCATTTATGAGGTTATATCAACTTATCGCACAAATTTG	501
1119	Db	TAAATTTATATTTTAAATATATATATATTAATATATATAATATATATTTTAAATATA	1060
502	Qy	TCATCGTAGTATATATTTCTGCAATTTCTCTTTCA'TTGAAGTATTTTATAGAAAAATTTAT	561
1059	Db	TTATTTATTAATATAGAATA'TTATAAAATATATAATTTTTTATATTTTTTTTTTATAATAAT	1000
562	Qy	TTAATATTTCTAGTATGAGAGATACAC'TGAAAGTTATTTGTCGTGTTGTTTGTTCGTTT	621
999	Db	TTATTTTATATTTATTTTTTTTTTATTTATATAAAATTAATATATTTTATATATAAAATATT	940
622	Qy	GT'TGTGTTGTTTTTCAAAATATAGAAAT'ACAAGAAATATAGTTGGCGCTTATTTTGCAA	681
939	Db	TTATTTATTTTAA'TTAA'TTAAATATATATATATATATATTTATGTTAA'TTTTATATA	880
682	Qy	TAAGTATTTGTTTAAAT	741
879	Db	TTATATTTATTTTATATTTTTTTTTTATAAATATAAAATA-----TTATATAATTTTTTAT	828
742	Qy	ATTACTCAC'TTGCA'TTTAAGAAAT'AAATGAGATATAACATGCTGGCTTAACAAATGATAG	801
827	Db	ATATTTTAA'TTTTATTTAA'TTATTAATATATATATTAATTTATTTTATTTAAATTTAT	768
802	Qy	ATAAA'TTCATATATCCATGTTATTTGTTTATTTGAA'GTGCTGTGTTTTTAAAGTTTAA	861
767	Db	ATATATTTAA'TTATPAAATTTTTTAAATAATTTATAATTAATATATATAAAATTTAA	708
862	Qy	GTTTCATAAAAGAAAGTATATAAACTTTCTTGGATTTTGTGACTTTATTTTATATGCTG	921
707	Db	AA'TTAAATTTTATATAATTTATTTAAATTTATTTTTTATATATTTTAAATATTTAAATTTA	648
922	Qy	TGATGATGATGTTTGTGTTGTCGAGCGTATACAAACAAATCCCATGAGCGTTTATTTTCA	981
647	Db	TTATATATATATA'TTTTATATATATTTTATTAATTTATTTTAAATTTTATTTTATAA	588
982	Qy	CTAACACTGTTTATCTCGATTAAAA'TCCAA'TTATTTATCGTTGGGTTATTTTATAGGAC	1041
587	Db	TAATTTTATATATATATATTTTTTATTTTAAATATATATATTTTAAATTTATAAAATATATAGT	528
1042	Qy	GAA'TGTTCCTTATG--AGATATTTTCTTTATCTCTAAATATTTCTCTTTATCATAAATCT-	1097
527	Db	TTAATAATTTTATTTATTTTAAATAATTTAAATTTAAATTTTTTATAATTTATATAAAATA	468
1098	Qy	-----AAGCCCTATGAATCGGCTCTCCGVAAATTTMARGAATATATTTTAAATTAATGG	1152
467	Db	TAATAAATATATTTAA'TTTTTTTTATTAATAAATTTATTTATATATTTTATTTATTA	408
1153	Qy	GGAAAGTCGACCTCAAGTCATTTTATGTCCTTTGGTTACGGAGAAAAACATCTCCCTT	1212
407	Db	TTATAATATATATTTTATTTATATAATAAAATATATATATATATAATAATAATATA	348
1213	Qy	TTTTTAGATGCTTAAATATATAATACAGGAGCTCTGTTGTTGTAACCTTATTCAGGAGGAA	1272
347	Db	ATT'TATATATAATTTTATATAATTTTGAATATATAATATGTATATATTTTATTTAATAT	288
1273	Qy	AGCTAACAGCAATTTCTTTACCAATGTTTTTTTTTAAATGCAATTCCTTAYCCAAATGGAATAC	1332
287	Db	AA'TTTATATGTTTATATATTTTATATTTTATAATAATTTAAATTTAAATTTAAATTTAA	228
1333	Qy	AACAGATAGCTAAGGAGATACGAATTTTATTTTAA'TTTA	1370
227	Db	TTTAAATAATAATTTTATGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA	190
RESULT 4			
CL078538/c			
LOCUS			
CL078538		1536 bp	DNA linear
		GSS	31-DEC-2003

RESULT	4
CL078538/c	
LOCUS	CL078538 1536 bp DNA linear GSS 31-DEC-2003
DEFINITION	CH215-15IC11 SP5.1 CH216 Xenopus tropicalis genomic clone

```

CH216-151C11, genomic survey sequence.
CL078538
CL078538.1 GI:40534451
GSS.
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 1536)
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp5 atctgcgcttcgacatc
Class: BAC ends
High quality sequence start: 1061
High quality sequence stop: 1145.
FEATURES
source
1..1536
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-151C11"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pPARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"
ORIGIN
Query Match 6.2%; Score 102.8; DB 9; Length 1536;
Best Local Similarity 46.2%; Pred. No. 4.6e-09;
Matches 346; Conservative 0; Mismatches 402; Indels 1; Gaps 1;

Qy 340 TCTGTGTAATTAATAGATAAAATATGTACAGGAGAGATGATTTCTCTAGGA 399
Db 1478 TTTTCTTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTNA 1419
Qy 400 TTTAGGATTTTGTAGGCAACGTTACGATGCTCTGCACATAAATAGAAATATTAT 459
Db 1418 TTTATTTTATTTTAAATTTTATTTTAAATTTTAAATTTTATTTTATTTTATTTT 1359
Qy 460 TTTGTTACTTTATGAGGTTATATCAACTTATGCACAAATTTGTCATCGTAGTATTAT 519
Db 1358 TTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1300
Qy 520 CTGCAATCTCTTCAATAGAGATTTATATAGAAAATTTATTAATTTCTAGTATG 579
Db 1299 TTTTCTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1240
Qy 580 AGAGATACACTGAAAGTATTGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTTCAA 639
Db 1239 ATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1180
Qy 640 AATATAGAAATPACAGAAATATTAGTGGCGCTTTATTGCAATAAGTATTGTTGTTAATA 699
Db 1179 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1120
Qy 700 ATGTACATATGACAGTATACCAATCTTGGATGACCTGTAATTAATCTACCTGCAATTA 759
Db 1119 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1060
Qy 760 AAGAAATTAATGAGATAACAAATCGCTGCTTAACAATGATAGATAAAATTCATATATCCAT 819
Db 1059 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1000

CH216-151C11, genomic survey sequence.
CL078538
CL078538.1 GI:40534451
GSS.
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 1536)
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp5 atctgcgcttcgacatc
Class: BAC ends
High quality sequence start: 1061
High quality sequence stop: 1145.
FEATURES
source
1..1536
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-151C11"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pPARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"
ORIGIN
Query Match 6.2%; Score 102.8; DB 9; Length 1536;
Best Local Similarity 46.2%; Pred. No. 4.6e-09;
Matches 346; Conservative 0; Mismatches 402; Indels 1; Gaps 1;

Qy 340 TCTGTGTAATTAATAGATAAAATATGTACAGGAGAGATGATTTCTCTAGGA 399
Db 1478 TTTTCTTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTNA 1419
Qy 400 TTTAGGATTTTGTAGGCAACGTTACGATGCTCTGCACATAAATAGAAATATTAT 459
Db 1418 TTTATTTTATTTTAAATTTTATTTTAAATTTTAAATTTTATTTTATTTTATTTT 1359
Qy 460 TTTGTTACTTTATGAGGTTATATCAACTTATGCACAAATTTGTCATCGTAGTATTAT 519
Db 1358 TTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1300
Qy 520 CTGCAATCTCTTCAATAGAGATTTATATAGAAAATTTATTAATTTCTAGTATG 579
Db 1299 TTTTCTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1240
Qy 580 AGAGATACACTGAAAGTATTGTCTGTTGTTGTTGTTGTTGTTGTTGTTTCAA 639
Db 1239 ATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1180
Qy 640 AATATAGAAATPACAGAAATATTAGTGGCGCTTTATTGCAATAAGTATTGTTGTTAATA 699
Db 1179 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1120
Qy 700 ATGTACATATGACAGTATACCAATCTTGGATGACCTGTAATTAATCTACCTGCAATTA 759
Db 1119 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1060
Qy 760 AAGAAATTAATGAGATAACAAATCGCTGCTTAACAATGATAGATAAAATTCATATATCCAT 819
Db 1059 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1000

CL143963 1205 bp DNA linear GSS 05-JAN-2004
ISBI-12302 Sp6.1 ISBI Xenopus tropicalis genomic clone ISBI-12302,
Genomic survey sequence.
CL143963
CL143963.1 GI:40637598
GSS.
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 1205)
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 600
High quality sequence stop: 664.
FEATURES
Location/Qualifiers
1..1205
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISBI-12302"
/clone_lib="ISBI"
/notes="Vector: pBelOBAC11; ISBI-1 xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match 6.2%; Score 102.6; DB 9; Length 1205;
Best Local Similarity 44.0%; Pred. No. 5e-09;
Matches 349; Conservative 2; Mismatches 440; Indels 2; Gaps 1;

Qy 429 ATTGCTCTGCACATAAATAGAAATTTATTTTGTCTTATGAGTTATATCAACTT 488
Db 1202 ATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1143
Qy 489 ATGCGCAATTTGTCTCGTAGTATTATATCTGCAATTTCTCTTTTATTAGAGTTAT 548
Db 1142 TTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1083
Qy 549 TATAGAAATTAATTAATTTTCTAGTATTGAGAGATACACTGAAAGTATTGTCGTT 608

```



```
ACCESSION      CL078589
VERSION        CL078589.1  GI:40534502
KEYWORDS       GSS.
SOURCE         Xenopus tropicalis (western clawed frog)
ORGANISM       Xenopus tropicalis
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
                Xenopodinae; Xenopus; Silurana.
REFERENCE      1 (bases 1 to 1493)
AUTHORS        Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
                Mardis,E. and Wilson,R.
TITLE          A physical map of the xenopus tropicalis genome
JOURNAL        Unpublished (2003)
COMMENT        Contact: Richard K Wilson
                Genome Sequencing Center
                Washington University School of Medicine
                Email: submissions@watson.wustl.edu
                Insert Length: 175000 Std Error: 0.00
                Seq primer: Sp5 atctgcgttcgatcct
                Class: BAC ends
                High quality sequence start: 1009
                High quality sequence stop: 1070.
FEATURES       source
                location/Qualifiers
                1..1493
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /strain="Nigerian frog"
                /db_xref="taxon:8364"
                /clone="CH216-151K5"
                /sex="male"
                /cell_line="Stock 248 F7A2, inbred N7"
                /clone_lib="CH216"
                /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                BAC library"
ORIGIN
Query Match      6.2%; Score 102.2; DB 9; Length 1493;
Best Local Similarity 46.6%; Pred. No. 6e-09;
Matches 327; Conservative 0; Mismatches 372; Indels 3; Gaps 1;
QY 387 TATTTCCTAGGATTAGGATTTGTTAGGCAACGTTTACGATGCTCGACAATAAT 446
DB 1142 TTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1083
QY 447 TAGAATATTATTTTCTTACTTATGAGGTATATCAACTATGCGACAATTTGTCATC 506
DB 1082 TTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 1023
QY 507 GTAGTATTATTCGCAATCTCTTTTCATAGAGTTAATGAGTATTTATAGAAAATTAAT 566
DB 1022 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 963
QY 567 ATTCTAGTAGTACACACTGGAAGTATTGTCGCTGTTGTTGTTGTTGTTGTTG 626
DB 962 TTTTATTTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 903
QY 627 TTTGTTTTTCAAAATATAGAAATACAAATATTAGTTGCGCTTTATTTGCAATAAGT 686
DB 902 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 846
QY 687 ATTGTTGTTAATGATGACATTAAGCAGATACCAATCTTTGGATTTGGACCTGTAATAC 746
DB 845 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 786
QY 747 TCACTTCGATTTAAAGAAATTAATGAGATAACAAATCTCGCTTAAACAATGATAGATA 806
DB 785 TTTATTTAATTAATAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 726
QY 807 TTCATATATCCATGTTTATTTGTTTATTTGAAGTTGCTGTGTTTAAAGTTTAAAGTTTC 866
DB 725 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 666
QY 867 ATAAAGAAAGATATAATAACTTCTTGGATTTTGACATTTATTTTATGCTGTGATG 926

Db 665 TAATAAAAAAATAAATTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 606
QY 927 ATGTATGTTTTTTCGAGCGGTATACACAATAATCCATGAGGTTTATTTTACCACTAAC 986
DB 605 TTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 546
QY 987 ACTGTTTATTCGATTAATAAATCCAAATTTATCGTGGTATTATTTATAGGACGAAT 1046
DB 545 NTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNT 486
QY 1047 GTTCCTTATGAGATATTTTCTTTATCTAATAATTCCTCTTTAT 1088
DB 485 NTNTTNTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNT 444

RESULT 8
CL080711
LOCUS
DEFINITION
ACCESSION
VERSION
SOURCE
ORGANISM
Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
AUTHORS
Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE
A physical map of the xenopus tropicalis genome
JOURNAL
Unpublished (2003)
COMMENT
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RM4 ctcaaggcgcatcggtcgagc
Class: BAC ends
High quality sequence start: 690
High quality sequence stop: 766.
FEATURES
source
location/Qualifiers
1..1805
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-15918"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"
ORIGIN
Query Match      6.2%; Score 102; DB 9; Length 1805;
Best Local Similarity 43.3%; Pred. No. 6.5e-09;
Matches 546; Conservative 4; Mismatches 706; Indels 4; Gaps 2;
QY 329 ATAGTTAGAAATCTTGTGTAATTTAATTAGATAAAAAATTTGTACAGGGTAGAATGTA 388
DB 466 ATATTATTATTATTATTATTGATTTTNTTTTATAATAATTTATTTTATTTTCTNTATGT 525
QY 389 TTTTCTAGGATTTAGGATTTGTTAGGCAACGTTTAGAGTTGCTGACATAAATTA 448
DB 526 TTTTATATTTTATATAATAATTTTATTTTATTTTATTTTATACATTAATTTATA 585
QY 449 GAATTATTATTTTGTGTTACTTTATGAGGTATATCAACTTATCGCAAAATTTGTCATCGT 508
DB 586 CATATTTACATTTATATTATTATTTTACATTTTATTTTATTTTATTTTATTTTATTT 645
```

QY	509	AGTATTATATCTGCAATCTCTTCATAGAGTTATTTATAGAAAAATTAATTAAT	568
Db	646	ATTTTTTATTTTTTATTTTTTATTTATTTATTTTTTATTTTTTTTTTTTTTT	705
QY	569	TTCTAGTATTGAGAGATACACGTAAAGTTATTTGCTGTTGTTTGTGTTTGT	628
Db	706	TT	765
QY	629	TGTTTTTCAAATATPAGAATTAACAAGATAATTTAGTTGGCGCTTATTTTGC	688
Db	766	TTTCTTTTTTTTTTTTTTATCTATATTTTTTTTTTTTTTTTTTTTTTTTTT	825
QY	689	TGTTGTTAATATGACATATGACAGTATACCAATCTTGGATGGACCTGTTAAT	748
Db	826	TTTTTTTATTTTTTTTTTTTTTTCATTTATTTTTTTTTTCTTTATTTTTTT	885
QY	749	ACTTGCAATTAAGAAATTAAGAGATAACAATGCTGGCTTAACAATGATAGATA	808
Db	886	TTATTTTGTCTTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTTT	945
QY	809	CATATATCCATTTGTTATTTGGTTATTTGAAGTTGCTGTGTTTTTAAGTTT	868
Db	946	TATTTATTTTAAATTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTATAT	1005
QY	869	AAAAAGAAAGTATATAAACTTCTTGGATTTTGGACTTATTTTATGCTGTGA	928
Db	1006	AATATTTTATTAATAATATTTCTTTTTTTTTTTTTTTTTTTTTTAAATTT	1065
QY	929	GATGTTTTTTGTTTCGAGCGTATACAAACAAATCCCATGAGCGTTTATTT	988
Db	1066	TTTTTTTCAATTTAATAAATATATATTTATCATATATCATGTTTATATAA	1125
QY	989	TGTTTATCTCGAATTAATCCAAATATTTATCGTTGGGTATTTTATAGGAC	1048
Db	1126	TTTTTATTTT--TTTAAATATTTTATTTTTCATTTTCATTTTATATATAT	1183
QY	1049	TCCCTATGAGATATTTCTTTATCTAAATATCTCTTTATCAATAATCTAAG	1108
Db	1184	TTTATTTTAAATCAAAATATTTTTTATTTATTTATTTAGTTTACAAATAT	1243
QY	1109	ATCGGCTCTCGVAAATTTMARGATATAAATTTAATTAATGCGGGAAGTGC	1168
Db	1244	TTATATTTTACATATTTTATACATAAATTTATTTATATTTTACATATTT	1303
QY	1169	AAGTCATTTTAGTCTTTTGGTTACGGGAGAAAAACATCTCTTTTATAGAT	1228
Db	1304	TCGNAATTTAATAATTTTATTAACATTTATTTACTTTTATTTTAAATAT	1363
QY	1229	ATATAAATCAGGAGCTCTGTTGGTAAATCTTATTCAGAGGAAAGCTTAAC	1288
Db	1364	TATTATATATTTCTATTTATTTTTTTTTTATTTATATATCTTATATAAT	1423
QY	1289	TTTACCAATGTTTTTAAATGCA--ATYCTTAYCCAAATGGAATACACAGAT	1346
Db	1424	TTTATCTATTTTTTTTTTTTATTTTCTTAATATTTTTTTTATTAATCAT	1483
QY	1347	GGAGATACGAATTTATTAATTTAGCGAAGACGAGCTTTTCAGACATATTT	1406
Db	1484	AAAAATGATTTTATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTATCT	1543
QY	1407	GCTCAAGCTAGGATGATATGATATGATCAATTTTTTTAGGAGGAGCTTGA	1466
Db	1544	TATTTATTTTAAATATTTTACATATTTTAAATATTAATAATTAATTTAT	1603
QY	1467	ATTCGTTTTCCAGATAATGAGGTTATCTTTTAAGAGATTCATGCTGATAT	1526
Db	1604	ATTTTTTTAATAATTAATGTTTGTAAATTTTATTTTATATAATATTTAT	1663
QY	1527	CTTCTGCTTTTTAAAAATTAATTTTAGATAATGTTTATCATTTTGTGTTT	1586
Db	1664	TTACATATTTTTTATTTCTTATCATTTTTTTTCTTTTTTATTTATTTCT	1723
RESULT 9			
CF238805/c			
LOCUS			
DEFINITION			
IMAGE:6995950 5', mRNA sequence.			
ACCESSION			
CF238805			
VERSION			
CF238805.1 GI:33442013			
KEYWORDS			
SOURCE			
ORGANISM			
Xenopus tropicalis (western clawed frog)			
Xenopus tropicalis			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;			
Xenopodinae; Xenopus; Silurana.			
1 (bases 1 to 1626)			
NIH-MGC http://mgc.nci.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Daniela S. Gerhard, Ph.D.			
Office of Cancer Genomics			
National Cancer Institute / NIH			
Bldg. 31 Rm10A07 Bethesda, MD 20892			
Email: cgapbs-r@mail.nih.gov			
Tissue Procurement: Robert M. Grainger			
CDNA Library Preparation: Life Technologies, Inc.			
CDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLNL14680 row: 9 column: 21			
High quality sequence start: 71			
High quality sequence stop: 316.			
Location/Qualifiers			
1. 1626			
/organism="Xenopus tropicalis"			
/mol_type="mRNA"			
/db_xref="taxon:8364"			
/clone="IMAGE:6995950"			
/tissue_type="neutula"			
/dev_stage="embryo, stages 14-19"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NICHG_XGC_Emb6"			
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;			
Cloned unidirectionally. Primer: Oligo dr. Average insert			
size 2.1 kb. Constructed by Invitrogen. Note: This is a			
Xenopus Gene Collection (XGC) library."			
ORIGIN			
Query Match			
Best Local Similarity 6.1%; Score 101.4; DB 6; Length 1626;			
Matches 498; Conservative 3; Mismatches 739; Indels 0; Gaps 0;			
QY	340	TCTTCTGGAATTAATTAGATAAAAAATATTCAGAGGTAGAGTGTATTTCTCTAG	399
Db	1611	TTTTTTTNAATTTTTTTTTTATATAATTTTTTTTATATATATTTATTTATTTA	1552
QY	400	TTTAGCATTTTCTTAGGGCAACGTTTACGATTGCTCTGACAATAAATAGAATTA	459
Db	1551	TTAATTTATTTAATNANANATNTTTTTTTTTTTTTTTTATTTATTTATTTT	1492
QY	460	TTTGTACTTTATGAGGTTTATATCAACTTATCGCAAAATTTGTCATCGTAGTAT	519
Db	1491	TATTTTTTTTTTTTTTTTANATTTTTTCTTTNATTTAAATCTCTAAATTTTTT	1432
QY	520	CTGCAATTTCTTTCATTAGAGTTATTTATAGAAAATTTAATATTTCTAGTATT	579
Db	1431	TAAATTTATATATTTATATTTATTTATTTTNNAAATTTTTTTTTTTTATTT	1372
QY	580	AGAGATACACTGAAAGTTATTTGTCGTTGTTGTTGTTGTTGTTGTTGTTTCAA	639
Db	1371	TTTTTTTTTTTTTNNANNANTNANTTTTTTTTTTTTATTTATTTTATTTANNA	1312
QY	640	AATATAGAAATTACAAGAATATTAGTTGGCGCTTTATTTGCAATAAGTATTGT	699

Db 839 ATTATATATTTAAATATAATTTAAATATATTTATATATATATATTTATTTATATATTTATTT 898
QY 813 TATCCATCTGTTATTTGGTTTATTTGAAGTTGCTGTTTAAAGTTTAAAGTTTCATAAAA 872
Db 899 TTTTAAATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 958
QY 873 AGAAAGTATATAAATCTTCTGGATTTTGAATTTTGAATTTTATTTTATGCTGTGATGATAT 932
Db 959 TTAATATATATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1018
QY 933 GTTTTGTTCGAGCGTATACACAAATCCCATGAGCGTTTATTT--TCACCTAACACTG 990
Db 1019 TTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1078
QY 991 TTTATTTCTCGATATAAATCCAAATTTATTTATCGTTGGGTTATTTATTTATGACGAATGTTTC 1050
Db 1079 ATTTATTTATATAATTTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1138
QY 1051 CTTATGAGATATTTCTTTATCTAATTTCTCTTTATCATAAATCTAAGCCTATGAAT 1110
Db 1139 ATTTAAATTTATANTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1198
QY 1111 CGGCTCTCGVAAATTTMARGAATATAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1170
Db 1199 ATTTATTTATNTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1258
QY 1171 GTTATTTATGCTTTTGTGTAGCGGAGAAACATCTCTCTTTTATGATAGCTTAAAT 1230
Db 1259 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1318
QY 1231 ATAAATCAGAGCTCTGTGTGTAATACTTATTCAGGAGAAAGCTTAACAGCAATTTCTT 1290
Db 1319 AATTNTAATATATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1378
QY 1291 TACCAATGTTTTTAAATGCAATTCCTTAY 1319
Db 1379 TTTNTATNTATATATATATAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT

RESULT 11
CL083840/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CL083840 1599 bp DNA linear GSS 05-JAN-2004
ISB1-2H14.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-2H14,
genomic survey sequence.
CL083840
CL083840.1 GI:40558745
GSS.
Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1599)
Kremitski,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 1069
High quality sequence stop: 1158.
Location/Qualifiers
1. .1599
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-2H14"

FEATURES
source

CL038406 1594 bp DNA linear GSS 31-DEC-2003
CH216-46A6 Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-46A6,
genomic survey sequence.
CL038406
CL038406.1 GI:40494319
GSS.
Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1594)
Kremitski,C., Carter,J., McPherson,J., Warren,W., Graves,T.,

/clone_lib="ISB1"
/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match
Best Local Similarity 6.0%; Score 99.8; DB 9; Length 1599;
Matches 326; Conservative 0; Mismatches 378; Indels 0; Gaps 0;
QY 383 ATTTGTTATTTCTAGGATTTAGGATTTTGTAGGCAACGTTTACGATTCCTGCAAT 442
Db 1396 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1337
QY 443 AAATTAGAAATTTATTTTGTACTTTATGAGGTTATATCAACTATGCGCAATTTGT 502
Db 1336 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1277
QY 503 CATCTAGTATTTATTTCTGCAATTTCTTCTTTCATTAGAGTATTTATAGAAATTTAT 562
Db 1276 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1217
QY 563 TAAATATTTCTAGTATTTAGAGATACACTGAAAGTTATTTGTCTGTTGTTGTTGTTGTTG 622
Db 1216 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1157
QY 623 TTTGTTGTTTTCAAAATATAGAAATATAGAAATATTTAGTTGGCGCTTTATTTGCAAT 682
Db 1156 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1097
QY 683 AAGTATTTGTTTAAATATGATACATTTGAGTATACCAATCTGAGTTGGACCTGTTAA 742
Db 1096 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1037
QY 743 TTACTCACCTGCATTTAAAGAAATTAAGAGATAACAAATGCTGGCTTAACAATGATAGA 802
Db 1036 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 977
QY 803 TAAATTCATATATCCATTTGTTATTTTGGTTTATTTGAAAGTTGCTGTGTTTAAAGTTTAA 862
Db 976 TTTTATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 917
QY 863 TTTCAATAAAGAAAGATATATAAACTTCTTGGATTTTGTGACTTTATTTTATGCTGT 922
Db 916 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 857
QY 923 GATGATGATGTTTGTTCGAGCGTATACACAAATCCCATGAGCGTTTATTTATTTATTT 982
Db 856 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 797
QY 983 TAACTGTTTATTTCTCGATTAATAAATCCAAATTTATTTATGTTGGTTATTTATAGGACG 1042
Db 796 TTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 737
QY 1043 AATGTTCTTATGAGATATTTTCTTTATCTAAATTTCTCTTT 1086
Db 736 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 693

RESULT 12
CL038406/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

```
Mardis, E. and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
           Genome Sequencing Center
           Washington University School of Medicine
           Email: submissions@watson.wustl.edu
           Insert Length: 175000 Std Error: 0.00
           Seq primer: Sp6 ATTAGTGACACTATAG
           Class: BAC ends
           High quality sequence start: 471
           High quality sequence stop: 522.
FEATURES   Location/Qualifiers
            1..1594
             /organism="Xenopus tropicalis"
             /mol_type="genomic DNA"
             /strain="Nigerian frog"
             /db_xref="taxon:8364"
             /clone="CH216-46A6"
             /sex="male"
             /cell_line="Stock 248 F7A2, inbred N7"
             /clone_lib="CH216"
             /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
             BAC library"
ORIGIN
Query Match      6.0%; Score 99; DB 9; Length 1594;
Best Local Similarity 45.8%; Pred. No. 2.4e-08;
Matches 327; Conservative 0; Mismatches 387; Indels 0; Gaps 0;
Qy 384 TTGTTATTTCTAGGATTTAGGATTTTGTAGGCAACGTTTACGATTGCTCTGACAATA 443
Db 1557 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1498
Qy 444 AATTAGAAATTTATTTTGTACTTTATGAGGTATATCAACTATGCGACAATTGTGC 503
Db 1497 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1438
Qy 504 ATCGTAGTATATATTCGCAATCTCTCTTCATTAGAACTTATTTATAGAAAATTTATTT 563
Db 1437 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1378
Qy 564 AATATTTCTAGTATGAGATACACCTGAAAGTTATTTGTCGTGTTTGTGTTGTTGT 623
Db 1377 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1318
Qy 624 TTGTTGTTTCTTCAAAATATAGAAATTAACAGAAATATAGTTGGCGCTTTATTTGCAATA 693
Db 1317 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1258
Qy 684 AGTATTCGTGTTAATAATGATATATGACAGTATACCAACTTGGATTGGACCTGTTAAT 743
Db 1257 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1198
Qy 744 TACTCACTTCGATTTAAAGAAATTAATGAGATACAAATGCTGGCTTTAAACAATGATAGAT 803
Db 1197 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1138
Qy 804 AAATTCATATATCACTGTTTATTTGGTTATTTGAAGTTCGCTGTTGTTTAAAGTTAAAGT 863
Db 1137 TTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1078
Qy 864 TTCAATAAAAGAAAGTATATAAATCTTCTTGGATTTTTCGACTTTTATTTTATGCTGTG 923
Db 1077 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1018
Qy 924 ATGATGATATGTTTGTTCGAGCGGTATACAACAAAATFCCCATGAGCGGTTTATTTTCACT 983
Db 1017 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 958
Qy 984 AACACGCTTTATCTCATTAATAAATCAATATTTATTCGTGGTATTTTATAGGACGA 1043
Db 957 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 898
```

```
Qy 1044 ATGTTCCTTATGAGATATTTTCTTTATCTAATATTCCTCTTTATCATATAAATCT 1097
Db 897 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 844

RESULT 13
LOCUS   CL082000
DEFINITION CH216-165P18_RM4.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CL082000
VERSION    CL082000.1 GI:40537913
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 1981)
AUTHORS   Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
           Genome Sequencing Center
           Washington University School of Medicine
           Email: submissions@watson.wustl.edu
           Insert Length: 175000 Std Error: 0.00
           Seq primer: RM4 ctcaaggcgatcggtcgagc
           Class: BAC ends
           High quality sequence start: 265
           High quality sequence stop: 497.
FEATURES   Location/Qualifiers
            1..1981
             /organism="Xenopus tropicalis"
             /mol_type="genomic DNA"
             /strain="Nigerian frog"
             /db_xref="taxon:8364"
             /clone="CH216-165P18"
             /sex="male"
             /cell_line="Stock 248 F7A2, inbred N7"
             /clone_lib="CH216"
             /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
             BAC library"
ORIGIN
Query Match      6.0%; Score 98.6; DB 9; Length 1981;
Best Local Similarity 46.2%; Pred. No. 2.8e-08;
Matches 326; Conservative 0; Mismatches 379; Indels 0; Gaps 0;
Qy 384 TTGTTATTTCTAGGATTTAGGATTTTGTAGGCAACGTTTACGATTGCTCTGACAATA 443
Db 351 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 410
Qy 444 AATTAGAAATTTATTTTGTGTTACTTTATGAGGTATATCAACTATGCGACAATTGTGC 503
Db 411 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 470
Qy 504 ATCGTAGTATATATTCGCAATTCCTCTTCATTAGAAATTTATTTATAGAAAATTTATTT 563
Db 471 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 530
Qy 564 AATATTTCTAGTATGAGATACACTGAAAGTTATTTGTCGTGTTGTTGTTGTTGTTCT 623
Db 531 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 590
Qy 624 TTGTTGTTTCTTCAAAATATAGAATTAACAAGAATATTAGTTGGCGCTTTATTTGCAATA 683
Db 591 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 650
Qy 684 AGTATGTTGTTAATAATGATACATTTATGAGTATACCAATCTTGGATTGGACCTGTTAAT 743
```


Job time : 5402 secs

```
FEATURES      High quality sequence stop: 1114.
Source         Location/Qualifiers
1..1632
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-167P5"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match      5.9%; Score 98.4; DB 9; Length 1632;
Best Local Similarity 45.7%; Pred.No.3e-08;
Matches 353; Conservative 0; Mismatches 418; Indels 1; Gaps 1;

QY 328 AATAGTTAGAAATCTTGTGAATTTAATTAGATAAAAAAATATTGTACAGGGTAGAATTGT 387
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1608 ATTATTTTATATATTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1549
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 388 ATTTTCCTAGGATTTAGGATTTTGTGTAGGCAACGTTTACGATTGCTCGACAAATAAT 447
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1548 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1489
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 AGAATTATTTTGTCTTCTATGAGGTTATATCAACTTATCGCACAAATTTGTCATCG 507
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1488 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1429
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 508 TAGTATTATATCTCGCAATCTCTTTCATTAGAGATTATTTATAGAAAAATTTAATA 567
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1428 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1369
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 568 TTCTAGTATGAGAGATACACGAAAGTATTGTGCTGTTGTT-TGTTTGTGTTTG 626
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1368 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1309
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 627 TTTGTTTTTCAAAATATAGAAATACAGAAATATTAGTTGGCGCTTTATTTGCAATAAGT 686
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1308 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1249
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 687 ATTGTTGTTAAATATGACATATGACATACCAATCTTGGATCGACCTGTTAATAATAC 746
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1248 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1189
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 747 TCACCTTGCATTTAAAGAAATTAATGAGATAACAAATCGCTGCTTAAACAATGATAGATAA 806
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1188 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1129
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 807 TTCATATATCCATTTGTTTGGTTTATTTGAGATTGCTGCTGTTTATTTAGTTTAACTTC 866
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1128 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1069
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 867 ATAAAAGAAAGATATAAACTTCTGTGATTTTGTGACTTTATTTTATTTATGCTGTGATG 926
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1068 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1009
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 927 ATGTAAGTTTGTGCGAGGTATACAAACAAATCCGATGAGCGTTTATTTCCACCTAAC 986
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1008 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 949
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 987 ACTGTTTATTTCTGATTAATAATCCAAATTTATTTATCGTTGGTTATTTTATAGGACGAAT 1046
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 948 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 889
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1047 GTTCCTTATGAGATATTTCTTTATCTAAATATTCCTCTTTATCATATAATCTA 1098
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 888 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATA 837
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```